UNIVERSITY OF LONDON THESIS

Degree PhD Year 2006 Name of Author GRIFFIN Sarah Joanne

Sensitivity to Interaural Timing Differences within High-Frequency Sounds

Sarah Joanne Griffin

Submitted for the Degree of
Doctor of Philosophy
University College London

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

I, Sarah Joanne Griffin, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Interaural Timing Differences (ITDs) are a cue for sound localisation. In response to low-frequency sounds, sensitivity to ITDs can be conveyed by the fine-structure of the sound waveform. In response to high-frequency sounds, sensitivity to ITDs can only be conveyed by the amplitude modulated envelope of the sound waveform. Sensitivity to ITDs within high-frequency sounds has classically been described as poorer than in response to low-frequency sounds. However, using a "transposed" sound stimulus, it has been shown that human sensitivity to ITDs in high-frequency sounds can be equivalent to sensitivity to ITDs in low-frequency sounds. In the present study, sensitivity to ITDs was investigated in the responses of neurons from the Inferior Colliculus of the guinea pig using transposed, and conventional, stimuli. A neural correlate of the improvement in sensitivity to ITDs provided by transposed tones was found. ITD-tuning functions had greater depths of modulation in response to transposed tones as compared to conventional stimuli, and neural discrimination thresholds for ITDs in transposed tones were similar to those obtained in response to low-frequency tones.

Neural coding of ITDs at low frequencies has been shown to depend on a neuron's frequency tuning. Therefore, the responses of neurons were examined for evidence of frequency-dependent tuning to ITDs in the envelope of high-frequency stimuli. The frequency-dependent ITD-tuning that was found contradicts a model of ITD coding proposed in 1948 by Jeffress. ITD-coding at high-frequencies, similarly to at low-frequencies, may use a population of neurons which are broadly tuned to ITDs.

It is suggested that sensitivity to ITDs in the envelope of high-frequency sounds is restricted both by peripheral processing and also by an upper f_m above which sensitivity to ITDs does not occur. For these reasons, the physiological relevance of sensitivity to ITDs in the envelope of high-frequency may be limited.

Acknowledgements

The physiological data presented in this thesis was collected in the laboratory of my PhD supervisor, Professor David McAlpine, at University College London (UCL). I acknowledge the input of Professor McAlpine, Dr. Neil Ingham and Professor Leslie Bernstein in the initial design of the experiments presented in Chapter 3 (Griffin *et al.* 2005). Data collected from neurons with low frequency characteristic frequencies were obtained from the data archive of the McAlpine laboratory, unless noted otherwise. Thanks to David, John, Neil, Susan and Torsten for enabling this. Thanks to Alan Palmer and Trevor Shackleton for allowing me to use their low-frequency JND data.

The psychophysical data presented in results Chapter 5 was collected whilst visiting the laboratory of Professor Leslie Bernstein in Connecticut, United States of America, courtesy of a Bogue fellowship from UCL. These data constitute preliminary data from 2 subjects. Subsequent to collecting this data and after I had returned to the United Kingdom, a full data set from 4 subjects was collected. It was not considered appropriate that these data be included in this thesis and therefore only the preliminary data are presented. I acknowledge the input of Professor Bernstein and Dr. Tom Buell in the design of these experiments. The modelling studies presented in Chapters 4 and 7 were implemented by altering Matlab code initially provided by Professor Bernstein and previously published in Bernstein & Trahiotis (2001).

Thanks to the McAlpine lab (Neil, Torsten, Miles, Nicol, John, Ade and Isabel) for providing companionship during this PhD work and to David for supervising me. I've greatly valued the support of this lab. Thanks to Les, Tino, Tom and Chantal for their discussions and help during my Bogue fellowship, for which I am also grateful. Also essential to this thesis work have been my proof-reading team (David, Heledd, Helen,

Isabel, Les, Miles, Roberta, Teresa and Torsten), and the Wellcome Trust who sponsored my PhD.

Thanks to Szonya and Martin for being great flatmates and to Helen and Freia for your continued friendship. For improving my climbing and helping me escape London, thanks to University College London Union Mountaineering Club, North London Mountaineering Club, Connecticut Climbers and Mountaineers, UCL expeditions' fund and David (for letting me escape). Thanks to John for coping with all my indecision. Your love and support are immeasurable to me.

Thanks to Mum, Dad, Paul and Ian. I still have a base at Malmsmead.

Sarah

Contents

List of Figures

List of Tables

List of Equations

1. Introduction

In the auditory system a map of sound frequency exists in the cochlea, and cochleotopic projections maintain this map in the central nervous system. In the visual system a map of visual space exists on the retina, and projections into the central nervous system are retinotopic, maintaining this map. Unlike in the visual system, there is no map of space in the peripheral auditory system. Therefore, the location of a sound source must be computed by the central nervous system. Since Lord Rayleigh (Strutt, 1907) realised that interaural time differences (ITDs) were a cue used to localise low-frequency tones, numerous investigators have studied the sensitivity to ITDs using psychophysical and physiological techniques. ITDs occur because of the difference in the travel time of sound from its source to each ear, and for a fixed interaural distance depend on the location of the sound source. The ability to detect ITDs on the order of 20 μs a represents temporal sensitivity in mammals that is unrivalled in other sensory systems (Blauert, 1997). The physiology and anatomy of the binaural (two ears) auditory system is specialised to enable such temporal resolution (Oertel, 1999; Trussell, 1999). Overviews are provided of the topics of sound, the psychophysics of sound localisation and the anatomy and physiology of the auditory system in Appendices II.1-II.3, because much of this background information can be found in textbooks. It is suggested that the reader who is not familiar with these topics should consult these Appendices.

According to the duplex theory of sound localisation, ITDs are a cue for localising low-frequency sounds (< 1.5 kHz) and interaural intensity differences (IIDs) are a cue for localising high-frequency sounds (Stevens & Newman, 1936). Sensitivity to ITDs in pure tones only occurs for frequencies < ~1.5 kHz (Blauert, 1997). Action potentials in auditory nerve fibres (ANFs) are phase-locked to the period of low-frequency tones (Weiss & Rose, 1998). This temporal structure is required in the coincidence detection process that underlies sensitivity to ITDs (Yin & Chan, 1990). For high-frequency tones (with periods that are small compared to the maximum ITD

that can be experienced) different sound source locations on the horizontal plane may produce the same interaural phase difference (IPD). In these cases, the IPD (equivalently the ITD in the ongoing portion of the waveform) is an ambiguous cue for localisation. At high frequencies, the auditory system is not sensitive to the fast fluctuations in the sound waveform and ITDs cannot be detected in pure tones. IIDs, in contrast, are thought to be a more prominent cue for localising high-frequency sounds because it is at high-frequencies (with wavelengths larger than the distance around the head) that sounds are most effectively shadowed by the head.

The duplex theory is, to a first approximation, reflected in the anatomy and physiology underlying sensitivity to ITDs and IIDs in mammals (Joris *et al.*, 1998; Tollin, 2003). The medial superior olive (MSO) has been considered to be specialised for detecting ITDs in low-frequency sounds (Joris *et al.*, 1998) and the LSO has been considered to be specialised for detecting IIDs in high-frequency sounds (Tollin, 2003). Electrophysiological recordings confirm that phase-locked action potentials converge from each cochlear nucleus (CN) onto coincidence detection neurons in the MSO , generating sensitivity to ITDs (Goldberg & Brown, 1969; Yin & Chan, 1990; Spitzer & Semple, 1995; Batra *et al.*, 1997b). Sensitivity to IIDs occurs in lateral superior olive (LSO) neurons, which receive excitatory ipsilateral and inhibitory contralateral inputs (Boudreau & Tsuchitani, 1968). The characteristic frequencies (CFs; see Section II.3.3) of MSO neurons are skewed towards low-frequencies whereas the CFs of LSO neurons are skewed towards high-frequencies (Goldberg & Brown, 1969; Guinan *et al.*, 1972).

The dichotomy suggested by the duplex theory is contradicted by psychophysical experiments that demonstrate sensitivity to ITDs conveyed by the envelopes of *high-frequency* complex sounds (Klumpp & Eady, 1956; Leakey *et al.*, 1958; David *et al.*, 1959; Henning, 1974a; McFadden & Pasanen, 1976; Henning, 1980; Henning & Ashton, 1981; Nuetzel & Hafter, 1981). Although sensitivity to ITDs cannot occur in response to high-frequency pure tones or to the fine-structure of a high-frequency complex sound, the amplitude modulated envelope of a high-frequency complex

sound can provide sufficient stimulus-locking to enable sensitivity to ITDs to be generated.

All of the experiments presented in this thesis investigate sensitivity to ITDs within the envelope of high-frequency sounds. From hereon, *high frequency* refers to sounds ≥ 2 kHz, and *low frequency* refers to frequencies at which sensitivity can occur to ITDs in pure tones. Natural sounds are replete with amplitude modulated structure (e.g. Rosen, 1992). Understanding the limitations and capabilities of the auditory system in detecting ITDs in high-frequency sounds is of interest to further understanding of sound localisation, and the auditory system in general. It also has potential applications for the design of cochlear implants, which only stimulate the high-frequency auditory system (Clark, 2003). The experiments described in this thesis investigate sensitivity to ITDs in high-frequency sounds using stimuli that are controlled for their frequency content (spectrum), and for the modulation rates of their envelopes. Three series of experiments were conducted and are introduced in the three Sections below.

1.1. A comparison of sensitivity to SAM, transposed and pure tones

Human sensitivity to changes in ITD conveyed by high-frequency stimuli has typically been found to be poorer (i.e. thresholds are higher) than that measured with low-frequency stimuli (Yost *et al.*, 1971; Jones & Williams, 1981; Bernstein & Trahiotis, 1982). The lateral extent of the intracranial image produced by an ITD conveyed by high-frequency complex stimuli is typically smaller than that conveyed by low-frequency stimuli (Bernstein & Trahiotis, 1985b; Trahiotis & Bernstein, 1986).

It has been suggested that the relatively poor sensitivity to ITDs at high-frequencies is due to limitations in the neural circuits that analyse envelope ITDs (Joris, 2003).

Sensitivity to ITDs in the envelope of high-frequency sounds can occur in the firing rates of neurons in both the MSO and the LSO (Yin & Chan, 1990; Joris & Yin, 1995). Neural responses of high-frequency LSO neurons have been shown to be relatively more modulated in response to changes in IID than ITD, when compared to neurons in the MSO (Caird & Klinke, 1983; Joris & Yin, 1995). Sensitivity to ITDs in the LSO has been considered a by-product of sensitivity to IIDs (Tollin, 2003). In the MSO, there are relatively few neurons with high-frequency CFs (high-CF neurons; Goldberg & Brown, 1969; Guinan *et al.*, 1972).

In contrast, Colburn & Equissaud (1976) hypothesised that the differences in ITD-sensitivity at low- and high-frequencies could be accounted for by differential effects of peripheral processing on low- and high-frequency stimuli. Peripheral processing is reflected in the firing rates of ANFs and refers to the processes that occur in the auditory periphery [including the basilar membrane, inner hair cells (IHCs) and auditory nerve fibres (ANFs)]. Firing of action potentials in ANFs in response to a low-frequency tone can be approximated by half-wave rectification of the tone, because IHC depolarisation only occurs as the basilar membrane moves in one direction (Section II.3.2). The ANF response is "phase-locked" to the low-frequency tone (Figure 1.1A). Firing of action potentials in ANFs in response to a high-frequency, sinusoidally amplitude modulated (SAM) tone can be approximated by half-wave rectification of the stimulus envelope. The modelled ANF response follows the sinusoidal envelope of a SAM tone (Figure 1.1B). Although this timing information supports sensitivity to ITDs in the envelope of a SAM tone, the firing pattern lacks the distinct "off periods" (where the firing probability is zero and produced by half-wave rectification) that characterise responses to low-frequency tones (compare Figures 1.1A & B). It was this difference that Colburn & Equissaud (1976) hypothesised could account for the difference in sensitivity to ITDs at low and high frequencies.

In human psychophysical experiments, Bernstein & Trahiotis (2002) demonstrated, using "transposed" tones, that ITD-sensitivity at high frequencies could be

comparable to ITD-sensitivity at low frequencies. Sensitivity to ITDs in transposed tones was also found to be higher than sensitivity to ITDs in SAM tones (Bernstein & Trahiotis, 2002). Transposed tones are generated by multiplying a half-wave rectified low-frequency tone with a high-frequency tone. A transposed tone has carrier frequency (f_c) equal to the frequency of the high-frequency tone and modulation rate (f_m) equal to the frequency of the low-frequency tone. Despite differences in the sound pressure waveforms, it is predicted that the probability of action potentials occurring in ANFs is similar for high-frequency transposed tones and low-frequency tones (compare Figures 1.1A & C).

Figure 1.1D shows the power spectra of exemplars of SAM and transposed tones. SAM tones are characterized by three frequency components, the carrier (f_c) and two "sidebands" [($f_c + f_m$) & ($f_c - f_m$)]. A transposed tone is spectrally complex when compared to a low-frequency tone or a SAM tone; transposed tones have additional sidebands spaced at multiples of $2 \times f_m$. The lower the frequency of the low-frequency tone, the wider the separation between the spectral components in the derived transposed tone. These differences between SAM and transposed tones are discussed further in Section 2.1.5.

To investigate whether a neural correlate of improved sensitivity to ITDs in transposed compared to SAM tones was present in the Inferior Colliculus (IC), extracellular recordings were made from single neurons of the guinea pig (*Cavia porcellus;* Griffin *et al.*, 2005). The data obtained are presented in Chapters 3 and 4. Neural responses reflected greater sensitivity to ITDs for transposed tones than for SAM tones. Consistent with the psychophysical findings of Bernstein & Trahiotis (2002), neural discrimination of ITDs within transposed tones was comparable to neural discrimination of ITDs within low-frequency tones. Also consistent with these, and other, psychophysical (Bernstein & Trahiotis, 1994; McFadden & Pasanen, 1976; Nuetzel & Hafter, 1981) and physiological (Joris & Yin, 1998) findings, ITD sensitivity in neurons was greatest at low modulation frequencies (< ~250 Hz) for both SAM and transposed tones. The results suggest that, for those low rates of

modulation, the central mechanisms that mediate sensitivity to envelope-based ITDs are essentially equivalent to those that mediate sensitivity to fine-structure-based ITDs.

1.2. The salience of onset-ITDs

The second series of experiments presented in this work relate to the salience of onset-ITDs. The difference in the time at which sound arrives at each ear drum is referred to as the onset-ITD. When sounds are presented over headphones, the onset-ITD can be different to the ITD in the ongoing portion of a sound waveform (ongoing-ITD; Figure 1.2). It is important to be clear what is meant by an onset-ITD (Zurek, 1993). Unless otherwise stated, an onset-ITD, in the present work, refers to the difference between the times of onset at the two ears. This could also be called a gating delay. It is standard practice when investigating human sensitivity to ITDs in the envelope of high-frequency sounds to attempt to remove the influence of sensitivity to the onset-ITD by setting the onset-ITD to 0 μs e.g. McFadden & Pasanen (1976), Hafter *et al.* (1979), Henning (1980) and Bernstein & Trahiotis (1982). Investigators do not want to measure any sensitivity to the onset-ITD *per se*, which might be due to detection of the period of monaural (in one ear) stimulation that occurs when the onset-ITD is not 0 μs. Stimuli are turned on by gradually increasing its intensity ("ramping on") to avoid producing low-frequency cues to the ITD (Bernstein & Trahiotis, 1982). Ramping on the stimulus also has the effect of weakening any onset-ITDs cues (Abel & Kunov, 1983). Despite these precautions, there are no previous studies investigating the salience of an onset-ITD for the discrimination of ITDs within a *high-frequency* stimulus.

Natural sounds contain both onset-ITDs and ongoing-ITDs, and it is of interest to question how salient an onset-ITD is, compared to an ongoing-ITD, for human sensitivity to ITDs. Previously, it has been shown that onset-ITDs, in low-frequency tones or low-frequency noise, contribute little to discrimination of ITDs when

compared to ongoing-ITDs (Tobias & Schubert, 1959; Buell *et al.*, 1991). Zurek (1993) hypothesised that onset-ITDs are more salient for discrimination of ITDs within high-frequency sounds than low-frequency sounds. However, Macpherson & Middlebrooks (2002) present data from experiments measuring human sound localisation performance, using high-frequency noise, which suggest that localisation performance is weakly influenced by onset-ITDs.

In Chapter 5, experiments are described which were designed to assess the salience of onset-ITDs for discrimination of ITDs within a high-frequency stimulus. Discrimination thresholds for ITDs were measured from human listeners using SAM tones that varied in the nature of their onset-ITDs. The results show, for two listeners, that onset-ITDs are less salient, for discrimination of ITDs, than ongoing-ITDs. It is suggested that the salience of onset-ITDs, for discrimination of ITDs, is comparable for both low- and high-frequency sounds.

1.3. Neural representation of ITDs in the envelope of high-frequency sounds

The third series of experiments presented in this work explores the neural representation of ITDs within the envelope of high-frequency sounds with reference to a model proposed by Jeffress in 1948. Jeffress (1948) described a model which explains how neurons are rendered sensitive to ITDs and how ITDs are encoded within a population of neurons. He suggested that an array of coincidence detector neurons exists (Figure 1.3A), in which each neuron codes for a different ITD by its maximum firing rate. Coincidence detector neurons receive bilateral excitatory inputs which are phase- or stimulus-locked in response to sound at the left or the right ears. The coincidence detector neuron fires maximally when the inputs from each ear arrive coincidentally. As a function of ITD, the firing rate of a coincidence detection neuron is modulated as phase-locked inputs from each ear move in and out of phase with each other. In response to a low-frequency pure tone, the firing rate of a

coincidence detector neuron will be minimal when the inputs are out of phase and will vary cyclically with a period equal to the period of the tone.

If there is an internal delay in the latency of arrival of excitation to a coincidence detector neuron in response to sound at the left and right ears, the neuron will fire maximally at an ITD of opposite sign to this. For example, if the latency of excitation arriving at a coincidence detector neuron in response to sound at the right ear is 100 μs less than the latency of excitation from the left ear, the internal delay is - 100 μs. This neuron will fire at the highest rate when the sound arrives at the left ear 100 μs before the right ear and its "peak-ITD" is +100 μs. Jeffress (1948) proposed that axons of different lengths (axonal delay lines) could create differences in the conduction times of excitation to a coincidence detector neuron (Figure 1.3A). The ladder-like arrangement of delay lines proposed creates a neural map in which ITDs are represented by the maximum firing rates of coincidence detector neurons.

In the mammal, neurons in the superior olivary complex (SOC) are the first neurons in the ascending auditory system to show modulation of firing rate with changes in ITD. The subject of how well the MSO (and mammalian auditory system) fits the Jeffress model has been reviewed by Joris *et al.* (1998) and, more recently, by McAlpine & Grothe (2003) and by Palmer (2004). This will also be reviewed below.

Anatomical and physiological evidence for a map of ITDs in the MSO

Anatomical evidence that *contralateral* inputs to the MSO could form delay lines in the cat has been described by Smith *et al.* (1993) and Beckius *et al.* (1999). Axon collaterals arrive first in the rostral part of the MSO and more distant collaterals innervate the caudal MSO. Beckius *et al.* (1999) also found, from measuring axon lengths, a gradient in the lengths of 4/6 axons providing *ipsilateral* input to the MSO, opposite in direction to that of the contralateral inputs. However, the gradient of these axon lengths was shallower than for contralateral inputs.

Yin & Chan (1990) report a tentative map of peak-ITDs obtained by examining the locations and peak-ITDs of MSO neurons with similar CFs and across different animals. This map of ITDs corresponds to that which would be predicted from the anatomy i.e. increasing ipsilateral delay along the rostral-caudal axis. The proposed map of ITDs in the MSO runs orthogonally to the tonotopic gradient of the nucleus which runs (from low- to high-frequency) along the dorsal-ventral axis. Although tentative in the mammal, good evidence exists for a Jeffress-type arrangement of delay lines in the avian auditory system. In the barn owl, *in vivo* recordings and anatomical evidence demonstrate delay lines in the axons of inputs to the nucleus laminaris (NL) from the nucleus magnocellularis (NM) and a map of peak-ITD in the dorsal-ventral axis of the NL (Sullivan & Konishi, 1986; Carr & Konishi, 1988; Carr & Konishi, 1990). In the chick, contralateral axons from the NM form delay lines across the medial-lateral axis of the NL but the ipsilateral axons do not (Young & Rubel, 1983; Overholt *et al.*, 1992).

Evidence for a map of space in auditory nuclei "downstream" of the MSO

A map of peak-ITDs has not been found in any other nucleus in the ascending mammalian auditory system. Projections from the MSO to the IC form synapses in a distributed manner along an iso-frequency lamina; the tonotopic gradient in the MSO is therefore maintained in the IC, but any systematic map of ITDs running orthogonally to the tonotopic gradient in the MSO is not maintained (Oliver *et al.*, 2003). In the central nucleus of the IC of the barn owl there is a map of ITD and frequency (Wagner *et al.*, 1987) and in the external IC there is a map of auditory space, including both horizontal and vertical dimensions (Knudsen & Konishi, 1978). In the mammalian brain, there is an example of a neural map of auditory space in the superior colliculus. The peak firing rates of neurons to positions in auditory space are aligned with the neural map of visual space (Palmer & King, 1982; Middlebrooks & Knudsen, 1984). However, the auditory space map is probably not mediated by sensitivity to ITDs (Campbell *et al.*, 2005).

Coincidence detection in the MSO

There is strong evidence that MSO principal neurons act as coincidence detectors. MSO neurons are phase-locked to pure tones (Goldberg & Brown, 1969; Moushegian *et al.*, 1975; Yin & Chan, 1990) and a neuron's peak-ITD can be predicted from the phases of their peak responses to tones presented at the ipsilateral and contralateral ear (Goldberg & Brown, 1969; Yin & Chan, 1990; Spitzer & Semple, 1995; Batra *et al.*, 1997b). Firing rates of MSO neurons vary cyclically with the same period as the interaural phase difference (IPD) of the tonal stimulus, and of the tone itself. As the frequency of the stimulating tone becomes higher, ITD-tuning functions (the firing rate of a neuron as a function of ITD) become narrower, reflecting the shorter cycle of the stimulating tone (Crow *et al.*, 1978; Yin & Chan, 1990). The principle of coincidence detection has also been tested using correlated and uncorrelated noise (Yin & Chan, 1990). As predicted, MSO neurons are only sensitive to ITDs in correlated noise. The response of a single MSO neuron, as a function of ITD, has been modelled as the cross-correlation (Section II.1) of the waveforms at each ear [although Batra & Yin (2004) demonstrate deviations from this in MSO responses].

The bipolar morphology of MSO neurons, with segregated ipsilateral and contralateral inputs, is considered to be a specialisation for coincidence detection (Agmon-Snir *et al.*, 1998). The membrane properties of MSO neurons are also considered to be specialised for coincidence detection and include the presence of a low-threshold potassium conductance which shortens the temporal integration window of the cell (Svirskis *et al.*, 2004; Magnusson *et al.*, 2005; Scott *et al.*, 2005).

Tuning to an ITD – phase-frequency relationship

If Jeffress' model is correct, then neurons in the MSO fire maximally at an ITD that is of the same magnitude, but opposite in sign, to their "internal" delay. Therefore, an MSO neuron should respond maximally to the same ITD (peak-ITD), irrespective of

the frequency of a pure tone stimulus. This is shown for a model response in Figures 1.4A & B. In Figure 1.4C the same relationship is shown, but instead of the peak-ITD the *peak-IPD* is plotted. The peak-IPD increases as the pure tone frequency increases because the peak-ITD becomes a larger proportion of the period of the tone. The slope of the (linear) relationship between peak-IPD and stimulus frequency is called the characteristic delay (CD) and the phase at which the line intercepts the y-axis is referred to as the characteristic phase (CP) (Rose *et al.*, 1966; Kuwada & Yin, 1983). The CP represents the phase difference between the inputs which causes maximum output. For a Jeffress-type neuron that detects coincidences of purely excitatory inputs, the CP should be 0. In this case, the CD represents the "internal delay" of a neuron and is equal to the peak-ITD (to which the neuron responds maximally).

The CD and CP of SOC neurons have been investigated by Yin & Chan (1990), Spitzer & Semple (1995) and Batra *et al.* (1997b). Batra *et al.* (1997b) report that ~40 % of neurons in the SOC have a relationship between frequency and peak-IPD with systematic deviations from a straight line. This is not consistent with Jeffress' model. The presence of neurons with CPs close to 0.5 cycles represents another departure from Jeffress' model (Batra *et al.*, 1997a). A CP of 0.5 is predicted from coincidence detection of phase-locked excitatory and inhibitory inputs, as occurs in the LSO (Joris & Yin, 1995; Batra *et al.*, 1997a). The maximum firing rate of LSO neurons occurs when inputs are out of phase. For some neurons, CPs are not close to either 0.0 or 0.5 cycles (Batra *et al.*, 1997a). Batra *et al.* (1997b) suggest that a well-timed inhibition could explain the departures from the CD/CP model that is observed in the SOC of the rabbit.

The role of inhibition in the MSO

Neural inhibition is not a feature of Jeffress' model. Nevertheless MSO principal neurons receive inhibitory glycinergic input from the medial nucleus of the trapezoid body [MNTB; and lateral nucleus of the trapezoid body (Grothe & Sanes, 1993; Grothe & Sanes, 1994; Smith, 1995; Kapfer *et al.*, 2002)]. Brand *et al.* (2002)

demonstrated that inhibition influences ITD-tuning in the gerbil MSO, by blocking glycinergic inhibition iontophoretically with strychnine. When inhibition was blocked, ITD-tuning was such that neurons' peak-ITDs were distributed around 0-μs ITD, whereas when the inhibition was present, peak-ITDs were distributed around ITDs that were equivalent to ~45° IPD. This suggests that the peak-ITDs of MSO neurons cannot (at least not exclusively) be explained by a "hard-wired" internal delay created by differences in axonal lengths.

Brand *et al.* (2002) propose a model for ITD-tuning in which fast, well-timed inhibition shifts the peak-ITD of MSO neurons away from 0-μs ITD by effectively delaying the response to excitation at the contralateral ear. *In vitro* electrophysiology suggests that the time-course of inhibition is fast in the gerbil (but not as fast as modelled in Brand *et al.* 2002; Magnusson *et al.*, 2005). A model for the mechanism by which inhibition changes a neuron's peak-ITD, and which has no requirement for temporal precision or fast kinetics, was proposed by Zhou *et al.* (2005). In this model, ITD-tuning is dependent on the asymmetric morphology of MSO neurons (Smith, 1995). However, an asymmetric morphology is not a consistent feature of MSO neurons in the gerbil (Scott *et al.*, 2005).

The range of peak-ITDs represented

Although there is little evidence for a map of space in the mammalian auditory system (see above), one might still expect ITDs to be coded by the peak-ITDs of ITD-sensitive neurons, consistent with Jeffress' model. In this case, neurons with similar CFs should have peak-ITDs that sample the range of ITDs that can be experienced. However, the peak-ITDs of neurons in the IC are often outside the physiological range of ITDs that the animal can experience, especially for mammals with small heads (Palmer *et al.*, 1990). This is not consistent with Jeffress' model and suggests that, in the mammal, ITDs are not coded directly by the peak-ITDs of ITD-sensitive neurons.

In the guinea pig, peak-ITDs have been observed to depend on CF such that for the lowest-frequency neurons, peak-ITDs tend to be furthest from 0-μs ITD (McAlpine *et al.*, 1996, 2001). Peak firing rates were found to be distributed around 45° IPD, i.e. the peak-IPDs were described as roughly constant across CF (McAlpine *et al.*, 2001). The CF dependence of ITD-tuning has also been reported for neurons in the gerbil MSO (Brand *et al.*, 2002; Seidl & Grothe, 2005) and the cat IC (Hancock & Delgutte, 2004). McAlpine *et al.* (2001) note the CF dependence of ITD-tuning results in the maximum slopes of ITD-tuning functions being close to 0-μs ITD, independent of CF. A large proportion of the dynamic range of a neuron's firing rate as a function of ITD, is therefore within the physiological range of ITDs that a guinea pig experiences. McAlpine *et al.* (2001) propose an opponent model of ITD-coding, consistent with their physiological data, comprising two broadly tuned channels (Figure 1.5.1C). One channel responds maximally to sound from the left and the other channel responds maximally to sound from the right. Using an information theoretic approach, Harper & McAlpine (2004) suggest that the most efficient ITD-coding strategy for mammals with small heads and low-frequency hearing comprises two broadly-tuned channels. This arises because the peaks of ITD-tuning functions are broad compared to the physiological range of ITDs that can be experienced and carry little information compared to their slopes.

Stecker *et al.* (2005b) propose a two channel "opponent neural population" code for sound localisation based on responses of neurons in the auditory cortex. This model generates sensitivity to spatial location that is robust to overall changes in stimulus level. Previous models of spatial coding in the auditory cortex have employed distributed models relying on artificial neural networks to recognise patterns of responses e.g. Furukawa & Middlebrooks (2002).

Summary of the evidence for Jeffress' model

Although Jeffress' 1948 model has proved to be productive in driving the investigation of sensitivity to ITDs, it does not accurately describe the representation of peak-ITDs in mammals. Neurons in the MSO and LSO operate as coincidence detectors and generate neural sensitivity to ITDs. However, the role of inhibition and the departures from the CD/CP model suggest that fixed axonal delay lines do not, alone, produce tuning to ITDs. In mammals, there is only weak evidence for a map of peak-ITDs in the MSO and no evidence in other auditory nuclei. Neuronal ITD-tuning is dependent on CF in a manner which positions the greater change in firing rates of ITD-tuning functions within the physiological range. Therefore it has been proposed that the slopes, and not the peaks, of ITD-tuning functions code for ITDs.

The results relating peak-ITDs to the CF of auditory neurons have all been obtained from neurons with low-CFs (McAlpine *et al.*, 1996; McAlpine *et al.*, 2001; Brand *et al.*, 2002; Hancock & Delgutte, 2004). In this work, it was hypothesised that a similar relationship exists in the responses of high-CF neurons. The experiments described in Chapters 6 and 7 investigate the range of peak-ITDs from neurons with high-CFs. The results provide evidence that peak-ITDs of high-CF neurons are dependent on the f_m at which they are sensitive to ITDs. It is suggested that, similar to low-frequencies, a Jeffress-type matrix of neurons with different peak-ITDs does not code for ITDs in the envelope of high-frequency sounds.

1.4. Summary

ITDs are a cue for sound localisation generated by differences in the travel time for a sound to reach the left and right ears. Neural sensitivity to ITDs arises in the auditory brainstem through the detection of coincident arrival of stimulus-locked inputs generated from the band-pass filtered sound at each ear. Sensitivity to ITDs can occur in response to the fine-structure of low-frequency sounds, including pure tones, and

in response to high-frequency sounds in which the sound envelope is modulated in amplitude. Human sensitivity to ITDs within the envelope of high-frequency sounds has been described as poorer than sensitivity to ITDs within the carrier of low-frequency sounds.

High-frequency transposed tones improve (i.e. lower) thresholds for detection of ITDs compared to SAM tones with the same f_c and f_m. For some f_ms, thresholds for detection of ITDs in response to transposed tones are equivalent to those in response to low-frequency tones. In the experiments described in Chapters 3 and 4, transposed tones were used to investigate the sensitivity to ITDs of high-CF neurons in the inferior colliculus of the guinea pig.

An ITD can occur at the onset (onset-ITD) and throughout the ongoing portion (ongoing-ITD) of a sound waveform. The salience of an onset-ITD for discrimination of ITDs in a high-frequency SAM tone was investigated in experiments described in Chapter 5.

The nature of the neural code for ITDs has recently been questioned after the long-standing acceptance of a model proposed by Jeffress (1948). ITD-tuning of low-CF neurons is dependent on the sound frequencies to which they respond. Experiments described in Chapters 6 and 7 investigate the neural representation of ITDs within a population of high-CF neurons.

The results are discussed at the end of each Chapter and a final, general discussion will draw together conclusions from all the data presented.

1.5. Figures and Tables for Chapter 1

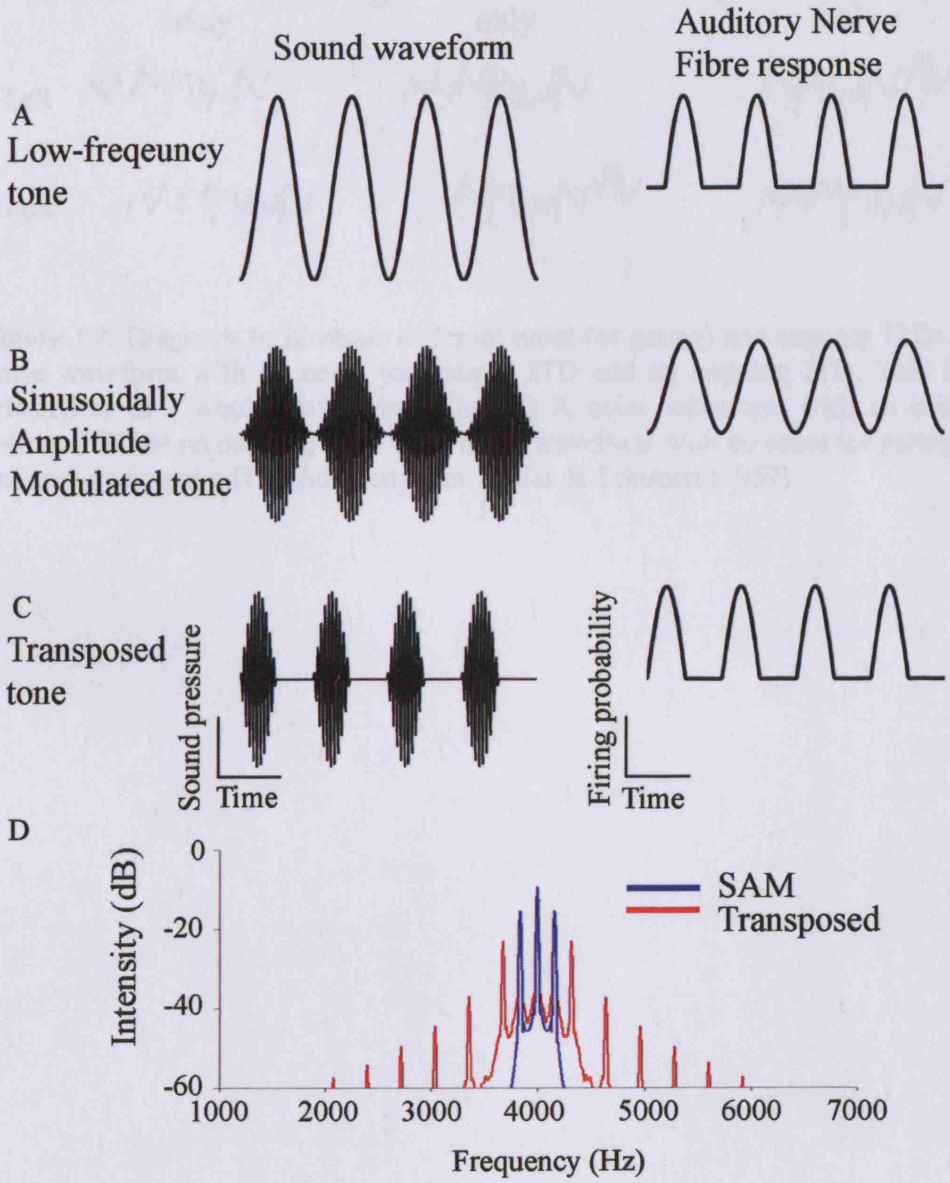

Figure 1.1. A diagram of a pure tone, a SAM tone and a transposed tone. Left: The sound pressure waveform of a low frequency tone (A), a SAM tone (B) and a transposed tone (C). Right: The results of peripheral auditory processing of each of the sounds (modelled by half-wave rectification and low-pass filtering). Adapted from Bernstein & Trahiotis (2002). D) The power spectra of exemplars of the SAM (blue) and transposed (red) tones used in the experiments described in Chapter 3. Spectra were recorded by presenting SAM (blue) and transposed (red) tones (f_c = 4 kHz and f_m = 160 Hz) to a spectrum analyser.

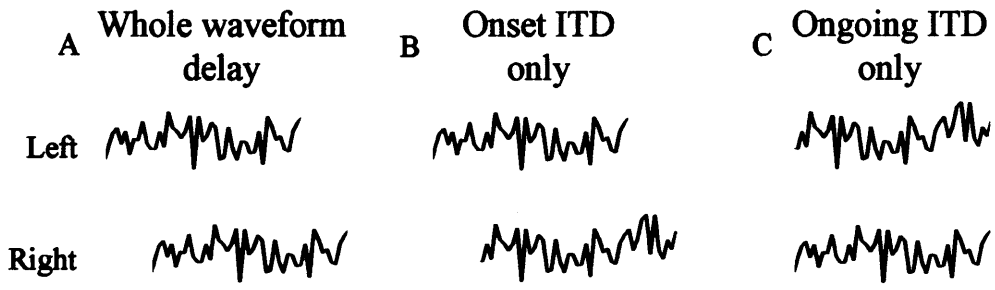

Figure 1.2. Diagrams to illustrate different onset (or gating) and ongoing ITDs. A) A noise waveform with an onset (or gating) ITD and an ongoing ITD. This is also referred to as a whole waveform delay. B) A noise waveform with an onset (or gating) ITD but no ongoing ITD. C) A noise waveform with no onset (or gating) ITD but with an ongoing ITD. Adapted from Tobias & Schubert (1959).

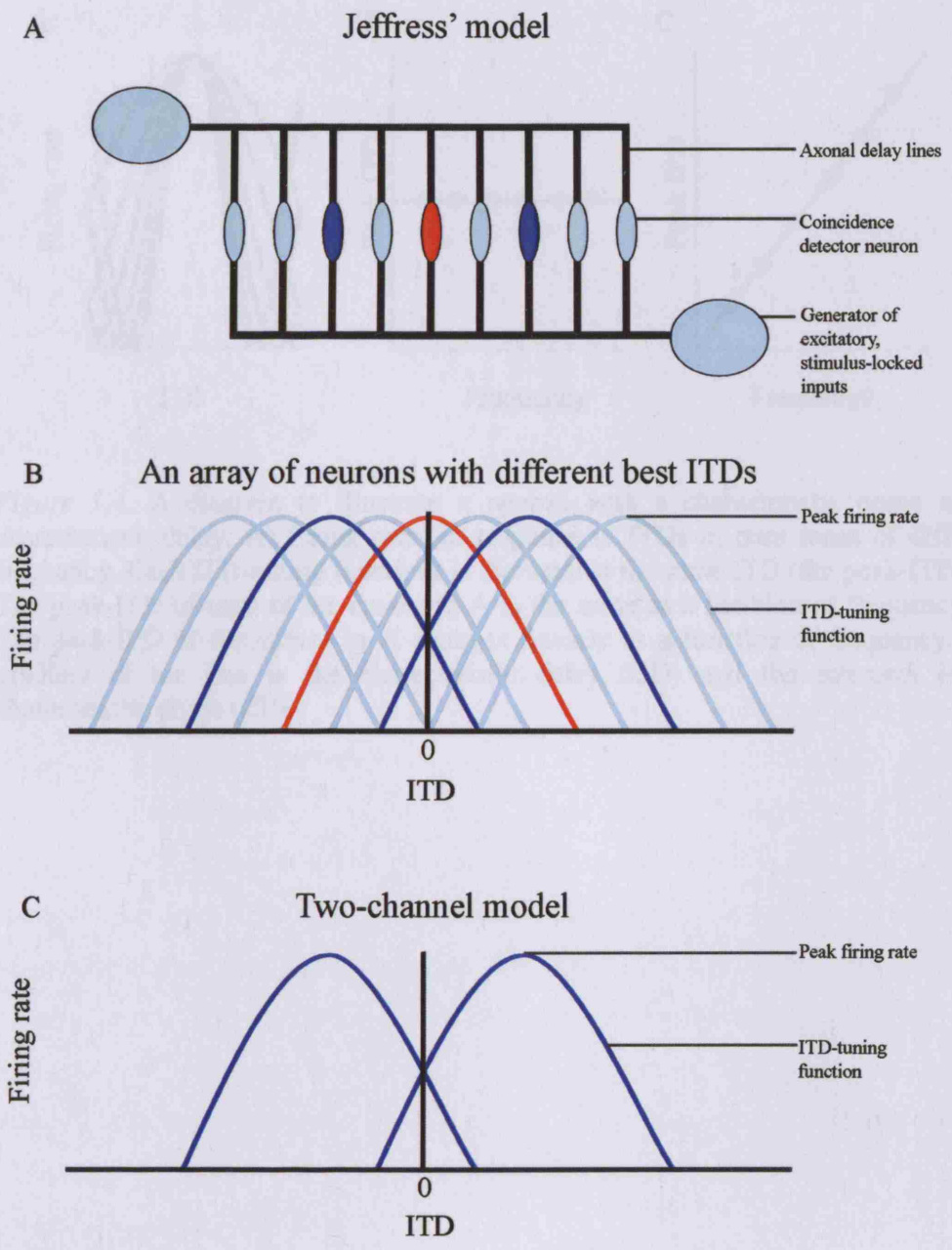

Figure 1.3. Models of ITD-coding. A) A representation of Jeffress (1948) model of ITD-coding. B) The firing rates of an array of neurons tuned to different ITDs as predicted by Jeffress' model. C) A representation of a two-channel model for ITD-coding. Only the neural firing rates around the peak ITD are shown in B and C.

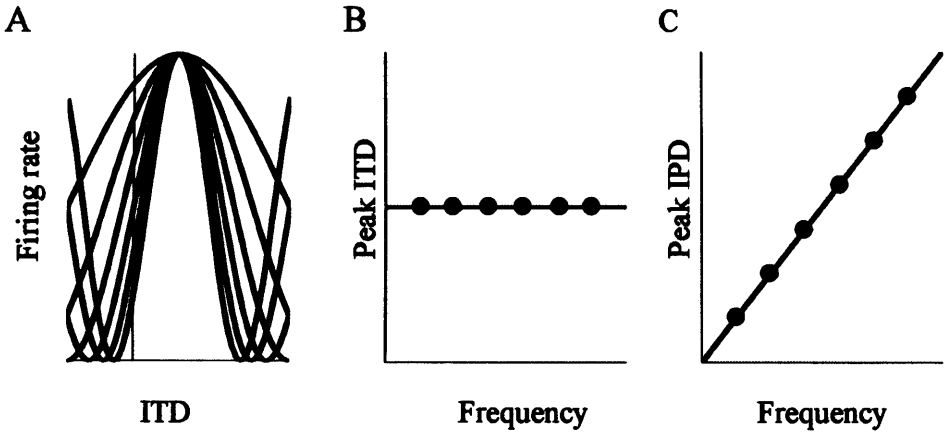

Figure 1.4. A diagram to illustrate a neuron with a characteristic phase and a characteristic delay. A) Firing rates in response to ITDs in pure tones of different frequency. Each ITD-tuning functions is maximal at the same ITD (the peak-ITD). B) The peak-ITD of each of the curves in A is the same as a function of frequency. C) The peak-IPD of the curves in A changes linearly as a function of frequency. The gradient of the line is the characteristic delay (CD) and the intercept is the characteristic phase (CP).

2. Materials and Methods

The materials and methods used in the work presented in this thesis are described in this Chapter in three Sections. These Sections describe the materials and methods used in the physiological experiments presented in Chapters 3, 4 and 6 (Section 2.1), the modelling presented in Chapters 4 and 7 (Section 2.2) and the psychophysical experiments presented in Chapter 5 (Section 2.3).

2.1. Physiology

2.1.1. Animals and surgery

All experiments were carried out in accordance with the Animal (Scientific Procedures) Act of 1986 of Great Britain and Northern Ireland. Single-neuron recordings were made from the central nucleus of the right IC of 63 adult guinea-pigs under urethane anaesthesia (Sigma-Aldrich, Poole, U.K.; 25% solution of 0.9% NaCl; 1g/kg, i.p.). Additional analgesia was provided using fentanyl-fluanisone (Hypnorm; Janssen-Cilag Ltd., High Wycome, U.K.; 0.1 ml, i.m.), supplementary doses of which were administered as required. Atropine sulphate (Animalcare Ltd., York, U.K.; 0.6 mg/ml; 0.1 mls, s.c.) was administered to reduce bronchial secretions, and lidocaine hydrochloride (Martindale Pharmaceuticals, Romford, U.K; 2%; s.c.) was administered locally prior to any surgical incision. A tracheal cannula was inserted. Core temperature was maintained at 37°C with a heating blanket and rectal probe (Harvard Apparatus Ltd, Kent, UK).

Animals were placed in a sound-attenuating chamber (IAC, Winchester, UK) and held in a stereotaxic frame with hollow ear speculae (modified from model 1730, David Kopf Instruments, CA). Before positioning the animal, the tragus was cut to obtain clear access to the tympanic membrane. A craniotomy was performed to expose the cortex overlying the IC, the covering dura was removed and agar (~2%)

was applied to prevent drying and deterioration of the cortex. The bullae were vented to equalize air pressure in the middle ears by insertion of cannulae and sealing with petroleum jelly. A parylene- or glass-coated tungsten microelectrode [1-5 MΩ; World Precision Instruments, FL; or manufactured in-house (Bullock *et al.*, 1988)] was positioned stereotaxically 2 mm above the IC (Medvedeva, 1977) and advanced ventrally from outside the recording chamber using a piezo-stepped microdrive (Burleigh Instruments, Inc., NY). At the end of each experiment, animals were administered a lethal dose of Sodium Pentobarbitone (Pentoject; Animalcare Ltd.; 60 mgs/ml; 1-2 mls, i.p. or Sagatal; 60 mg/ml).

2.1.2. Stimulus production and presentation

Sounds were produced using Tucker Davis Technologies (TDT; Alachua, FL) digital signal processing hardware. Tonal stimuli used for the isolation and characterization of single neurons were generated using custom software (Trevor Shackleton and Alan Palmer, MRC Institute of Hearing Research, Nottingham, UK; 100-kHz sampling rate) and TDT system II hardware. In later experiments, isolation and characterization stimuli were generated using RPvds and presented using TDT system III hardware. TDT Brainware, Real Time Processor Visual Design Studio (RPvds) and system III hardware were employed in order to generate the SAM and transposed stimuli (50-kHz sampling rate). A diagram illustrating the hardware used for stimulus production and presentation is shown in Figure 2.1A for the isolation and characterization stimulus and in Figure 2.1B for the SAM and transposed stimuli. The hardware used for all experiments is represented in grey, hardware used for production and presentation of the isolation and characterization stimuli is represented in blue and hardware used for the production and presentation of SAM and transposed stimuli is represented in pink.

Stimuli were generated and scaled such that their peak voltages were at the maximum voltage (± 10V) of the digital-to-analogue converters (DACs). The outputs were

attenuated to achieve the desired level for the experiments using PA4 (system II) or PA5 (system III) modules (TDT). (In initial experiments using system III, the attenuation was achieved via scaling within the signal-generation software.) Fixed amplification [ROTEL (Worthing, UK) RB971 power amplifier or a Beyerdynamic (Burgess Hill, UK) A150 Blueprint stereo-amplifier] was followed by 60 dB of "final" attenuation. Employing a high-level signal followed by attenuation provided a greater signal-to-noise ratio than if amplification alone were used to achieve the desired intensity.

Sounds were delivered via Beyerdynamic DT48A loudspeaker drivers, modified and fitted with a probe tube attachment to allow insertion and sealing into the hollow ear speculae. Knowles Acoustics (Burgess Hill, UK) FG3452 microphones (the "probe microphones") attached to steel tubes placed in the ear speculae and calibrated against a Bruel and Kjaer (Stevenage, UK) Type 4136 ⅛-in. microphone] were used to measure the stimulus within a few mm of the tympanic membrane in order to ensure that the sounds delivered to each ear were well matched. Sounds measured in this way were monitored via a spectrum analyzer throughout the experiments. System II hardware (shown in green in Figure 2.1A) could be used to store a digital record of the sound, which was analysed using Matlab (The Mathworks Ltd., Natick, MA). Maximum output was ~110 dB SPL, measured at 1 kHz.

2.1.3. Spike recording

Electrical signals from the electrode were conducted via a headstage to a pre-amplifier (TDT Medusa RA16PA) where they were amplified and digitized at a 25-kHz sampling rate. The signal was then conducted via a fibre-optic cable to the RA16 base station which produced fixed amplification and filtering (×1000 gain, 300-Hz high-pass filter, 10-kHz low-pass filter and 50-Hz notch filter). The hardware used for spike recording is represented diagrammatically in Figure 2.2. The hardware represented by grey boxes was common to all experiments. The hardware used

specifically for recording of spikes in response to isolation and characterization stimuli is represented in blue. Using system II hardware, a spike voltage – level discriminator (TDT ET1) was used to detect spikes from the background noise. Spikes were monitored on a Tektronix TD210 oscilloscope for continuity of spike characteristics, i.e., shape and amplitude, to ensure recordings were from a single neuron. The hardware used specifically for recording of spikes in response to SAM and transposed stimuli is represented in blue in Figure 2.2. Using system III hardware, spike data passed from the RA16 base station to TDT Brainware and spikes that crossed a user defined trigger level were counted. Spikes were then additionally sorted, in Brainware, according to characteristics such as shape and amplitude, to ensure data were obtained from a single neuron.

2.1.4. Sound stimuli - Isolation and characterisation stimuli

Binaural "search stimuli" consisting of diotic pure tones which were 50-ms in duration and repeated at 2-5 Hz to were used to isolate neurons and estimate their threshold (the lowest intensity sound required to evoke firing, determined audio-visually, by the experimenter) and characteristic frequency (CF; the frequency at which the lowest threshold was obtained). The CF and threshold were confirmed by recording a frequency-versus-level response area spectrally flanking the CF estimate (2 octaves above and 4 octaves below CF) and between 10 and 90 dB of attenuation from the maximum system output. If the neuron remained isolated for a sufficient length of time, neurons were further characterised by presentation of 50- or 200-ms diotic, ipsilateral and contralateral tones with frequency equal to CF. Tones were presented at 20 dB above threshold for 100 or 150 repeats.

2.1.5. Sound stimuli - SAM and transposed tones

SAM and transposed tones were presented in experiments described in Chapters 3 and 4. SAM and transposed stimuli were constructed with a carrier frequency (f_c; in

sine phase) equal to neuronal CF and were modulated by multiplication with either a lower frequency, dc-shifted, sinusoidal waveform (in sine phase) to produce a SAM tone (Figure 1.6.1B) or a half-rectified sinusoid (90° phase-advanced with respect to the SAM modulating tone) to produce a transposed tone (Figure 1.6.1C). The modulation depth for both stimuli was 100%. For each neuron, the peak voltages of SAM and transposed tones presented were equal and between 10 to 30 dB above the peak voltage of the sinusoid at CF that defined threshold. For 75/82 neurons, peak voltages were 20 dB above that of the tone at threshold. This choice resulted in the SAM and transposed tones being presented at 15.7 and 14.0 dB, respectively, above the *rms level* of the tone. In all cases, the rms levels of the transposed tones were within 2 dB of their SAM counterparts. Stimuli were 500 ms or 800 ms in duration and were gated on and off with 2-ms cosine-squared ramps. Two parameters of these stimuli were varied, ITD (of primary interest) and frequency of modulation (f_m).

ITDs (or IPDs)

An ITD was created by delaying the entire waveform (both fine-structure and envelope) in one channel and advancing it in the other by an equal amount. This resulted in an onset-, ongoing-, and offset-ITD in the stimulus (Figure 1.2.3). In early experiments the stimuli were ramped on and off at the same time in each ear; there was no onset or offset ITD. A positive ITD was created by delaying the stimulus in the ipsilateral (left) ear and advancing the stimulus in the contralateral ear. A negative ITD was created by delaying the stimulus in the contralateral ear and advancing the stimulus in the ipsilateral ear. The values of ITD that were presented depended on the f_m of the stimuli and are therefore referred to as IPDs. Seventeen IPDs were presented over ± 0.5 cycles of the f_m. The IPDs applied were the same irrespective of the f_m ensuring that the neural tuning was measured over a complete cycle of each f_m.

Modulation rate

Neural responses from one neuron, to any one f_m and at 17 IPDs will be referred to as a single "recording". SAM and transposed tones were both presented at each f_m. Initially, for each neuron, six or seven f_m were presented pseudo-randomly at each IPD. Modulation frequencies varied between 10 Hz and 640 Hz (or CF/6 if this was lower) in logarithmic steps or between 10 Hz and 650 Hz (or CF/6) in linear steps. For 11 neurons a smaller range of f_m was presented (between 10 Hz and 200-250 Hz) because it was found that sensitivity to ITDs was most likely to occur in this range. At least three (3-6, median = 3) repetitions of each stimulus condition were presented. The exact f_ms presented were determined by the experimenter according to the response of the neuron under investigation. For example, if the neuron remained well-isolated, further data sets were obtained at more narrowly spaced f_m around the f_m at which sensitivity to ITDs occurred.

Spectral components and the f_m limitation

van de Par and Kohlrausch (1997) demonstrated, in psychophysical experiments, that the five central frequency components (Figure 1.1.1D) of a high-frequency transposed stimulus were sufficient to yield improved binaural performance at high frequencies. For transmission of the temporal structure of the f_m to a single ANF, the spectral components of SAM tones and the central five spectral components of transposed tones must fall within the ANF's effective "filter" or spectral receptive field. As the f_m increases, sidebands are attenuated by the ANF's filter in proportion to their spectral "distance" from the CF. Accordingly, the f_m in the present study was limited to a maximum of CF/6 to include f_ms at, and slightly above, the range over which the temporal structure was expected to be preserved. Filter characteristics were judged from the shape of the frequency-versus-level response areas obtained from each neuron, and from reports of the equivalent rectangular bandwidth (ERB) of guinea pig ANFs (Evans *et al.*, 1992). The maximum f_m presented was limited to 640 or 650 Hz across all CFs because this represents the upper limit in the f_m at which

sensitivity to ITDs has been observed at high-frequencies (e.g. Bernstein & Trahiotis, 1994; Joris & Yin, 1998).

In their psychophysical experiments, Bernstein & Trahiotis (2002) limited the spectra of transposed stimuli to prevent the use of energy outside the psychophysically-determined "auditory filter" (see Moore, 1997) surrounding the centre frequency of the signal under investigation. Their half-wave rectified tones were first filtered to remove frequencies > 2 kHz before multiplication with a high-frequency carrier. It was not necessary to impose such a spectral limitation in the present study because the responses of single IC neurons were investigated. The spectral components influencing single neurons' responses are necessarily limited by their spectral receptive fields. Spectral components distant from f_c (outside the central 5 components) are unlikely to contribute any appreciable energy to a neuron with a CF equal to f_c and at the intensities at which transposed tones were presented (10 to 30 dB above pure tone threshold).

High-resolution functions

"High-resolution" IPD functions were obtained from 14 neurons. A f_m was selected to which a given neuron exhibited a substantial modulation in its firing rate as a function of IPD (i.e., was found to be IPD-sensitive). SAM and transposed tones with 101 (or 102) IPDs between +0.5 and -0.5 cycles, including an onset and offset ITD, were presented at this f_m. The duration of the stimulus was 500 ms and at least 6 repetitions (6 – 13, median = 10) were presented, pseudo-randomly, of each stimulus condition.

2.1.6. Sound stimuli – Transposed noise

Transposed noise stimuli were presented in experiments described in Chapter 6. Transposed noise was constructed by multiplying a high-frequency carrier tone by a flat-spectrum, low-frequency noise, referred to as the modulating noise, which was

first half-wave rectified (Figure 6.3). The f_c was equal to the CF of the neuron under investigation. The modulating noise was created digitally in the spectral domain with constant amplitude and random phase at each frequency component. For the data described in Sections 6.3.1, 2 & 4, the spectrum of the modulating noise only contained components below CF/5 (see Section 6.3.1 and Chapter 7). In the data described in Section 6.3.3 the low-pass cut-off of the modulating noise was varied. Stimuli were gated on and off with a 2 ms $\cos^2$ ramp.

Stimulus intensity

Correlated (and in later experiments also uncorrelated) transposed noise was presented over a 50 - 70 dB range of intensities in 10 dB steps to obtain functions relating firing rate to intensity. The range presented was dependent on the threshold of the neuron such that the function relating firing rate to intensity encompassed the dynamic range of the neuron (where the firing rate changed rapidly as a function of intensity). The transposed noise was then presented at an intensity either in the dynamic portion of the rate-intensity function or at an intensity at which firing rates in response to correlated and uncorrelated noise were different.

ITDs

Transposed noise stimuli were presented to each neuron with ITDs in steps of 200 μs over at least ± 2 ms. ITDs were most commonly presented between ± 4 ms or ± 8 ms. Other ITD ranges and ITD steps were used to explore further the shape of the ITD-tuning functions. ITDs were generated by introducing a linear phase delay in the modulating noise. The carrier tone was diotic. In initial experiments, the carrier was in sine phase but this was randomised in later experiments. 10-20 repeats were recorded at each ITD. If the isolation of the neuron deteriorated, the recording was terminated early resulting in four ITD-tuning functions with between 5 and 9 repeats.

Stimuli were either 200- or 1000-ms in duration and were repeated 300-ms or every 1600-ms. ITDs were presented in pseudo-random order.

Monaural stimuli

For 22 neurons, correlated transposed noise with the low-pass cut-off equal to CF/5 was presented to both ears and to each ear separately. A 200 ms duration was used and at least 6 repeats were presented.

2.1.7. Data analysis

Spike times were exported into Matlab 6.5 for off-line analysis. The data analyses performed are described in the Chapters in which the data are described.

2.2. Modelling

All modelling studies were implementing using Matlab.

2.2.1. Modelling – Input sound waveforms

Tones

410 ms samples of SAM and transposed tones (with random starting phase; length chosen for computational efficiency) were generated with initial amplitudes between ± 1 and a sampling rate of 40 kHz. The f_c was 4 kHz and the f_m was 80 Hz or 160 Hz. ITDs were introduced by delaying the modulating envelope (before half-wave rectification) of the left or right channel. Waveforms therefore had an ITD in their envelope but not in their carrier.

Noise

Noise stimuli were constructed in the frequency domain and each frequency had the same amplitude but random starting phase. In the time domain, they had a sampling rate of 50 kHz. The band widths were either 1.6 kHz centred on 4 kHz or 800 Hz centred on 500 Hz. The transposed noise was created as described in Section 2.1.6 and had a f_c of 4 kHz and a modulating noise with a low-pass cut-off of 800 Hz. All waveforms were 655 ms in duration (chosen for computational efficiency). An ITD was introduced using the same methods as for tones.

2.2.2. Modelling – Model of peripheral processing

The model described here is that developed by Bernstein & Trahiotis (2002) to account for the processing that occurs in the auditory periphery (see Section II.3.2-3). The sound waveforms were filtered with a gamma-tone filter to model the frequency-intensity tuning of an ANF (Patterson, 1994). The gamma-tone filter was implemented using the binaural toolbox created by Mike Akeroyd. The compressive effect of the basilar membrane was modelled by raising the waveforms to an exponent of 0.23. Half-wave rectification (square-law) was used to capture the rectification of the stimulus caused by the membrane potential of IHCs (Muller *et al.*, 1991). The waveforms were then filtered to stimulate the loss of phase-locking to the carrier using a Weiss/Rose filter with a corner frequency of 425 Hz (Weiss & Rose, 1988).

Bernstein & Trahiotis (2002) suggested the existence of another limit on the f_ms at which ITD-sensitivity occurs. A 150-Hz low-pass, second order Butterworth filter was included in their model to simulate the upper limit in the f_m at which ITD sensitivity occurs. This was similar to filters in models that account for *monaural* temporal modulation transfer functions (Ewert & Dau, 2000; Kohlrausch *et al.*, 2000).

2.2.3. Modelling - Coincidence detection and correlation

The waveforms, after modelled peripheral processing, were first shortened to remove the start and end of the waveform. The SAM and transposed tones were then 300 ms in duration and the noise stimuli were 635 ms in duration. Either the cross-product (Eq. 5) or the normalised correlation (Eq. 4) was calculated (see Section II.1.2).

2.3. Psychophysics

2.3.1. Stimulus production and presentation

Sounds were produced using Tucker Davis Technologies (TDT; Alachua, FL) system II, digital signal processing hardware using custom software (MLSig; Marcel van der Heijden). A diagram of the hardware used is shown in Figure 2.3.

All stimuli were generated digitally at either 50 kHz or, in later experiments, at 100 kHz (TDT AP2). Stimuli were then low-pass filtered at 8.5 kHz (TDT FLT2), attenuated (TDT PA4) to 70 dB SPL, measured from the output of TDH-39 earphones in a 6-cc coupler. A line amplifier was employed to increase the dynamic range of intensities that could be set on the TDT PA4. A low-pass (1.3 kHz cutoff and 30 dB SPL spectrum level) Gaussian noise [generated using Adobe Audition 1.5 and saved to CD] was presented continuously during the experiments from a CD player (TASCAN CD-RW 750).

All stimuli were gated on with a 10 ms $\cos^2$ ramp.

2.4. Figures and Tables for Chapter 2

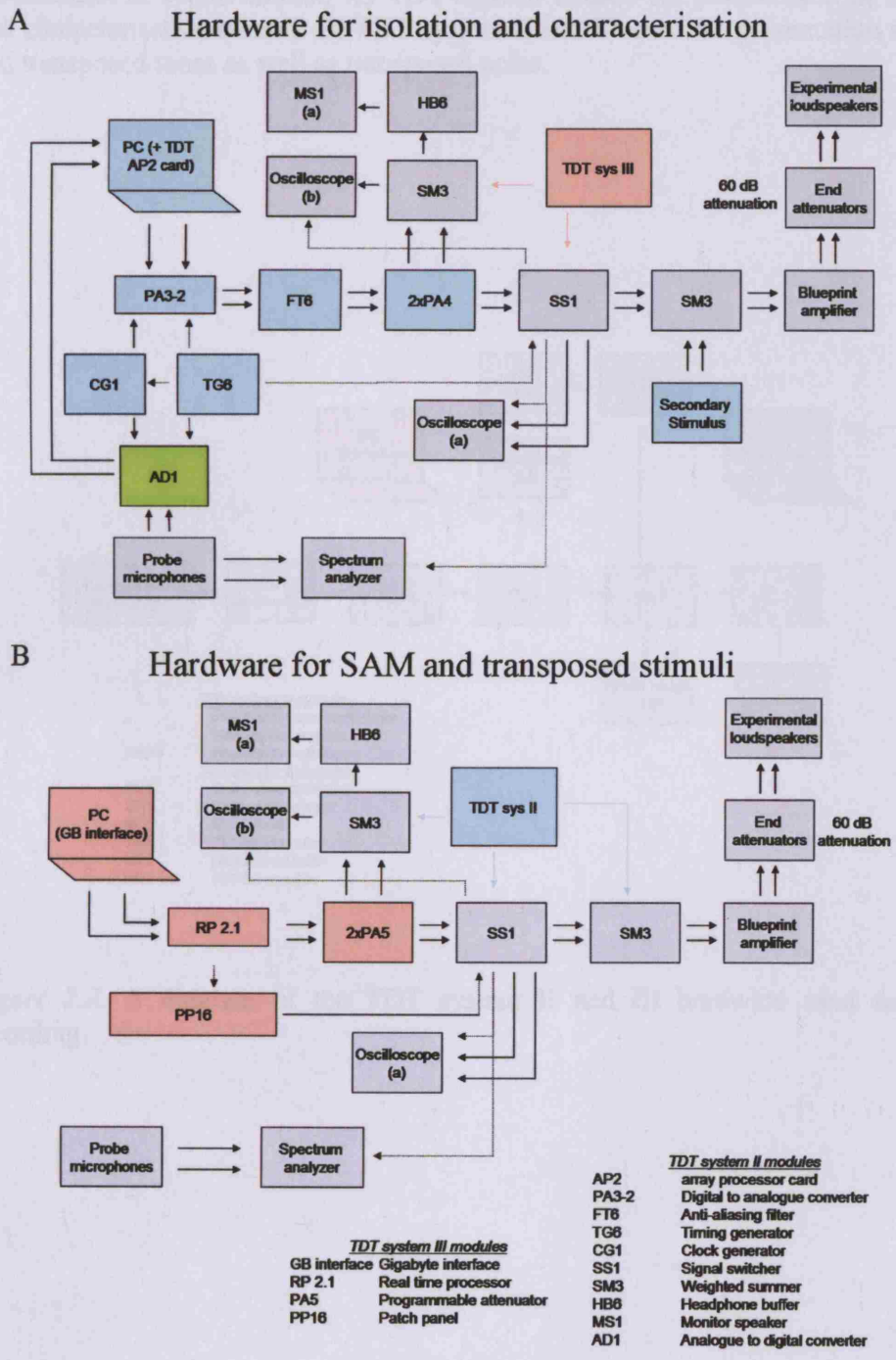

Figure 2.1. See following page for legend.

Figure 2.1. A diagram of the TDT system II and system III hardware used for the presentation of sound stimuli. A) TDT System II used for presentation of isolation and characterisation stimuli. B) TDT system II and III used for presentation of SAM and transposed tones as well as transposed noise.

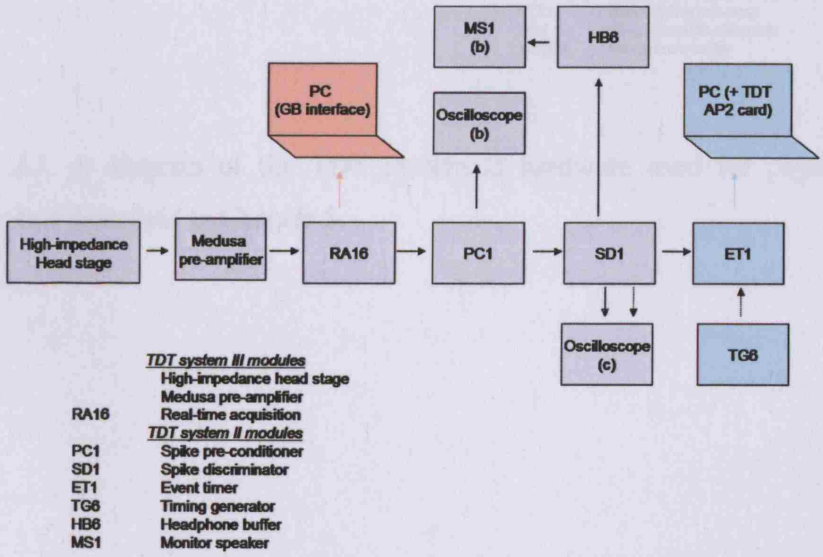

Figure 2.2. A diagram of the TDT system II and III hardware used for spike recording.

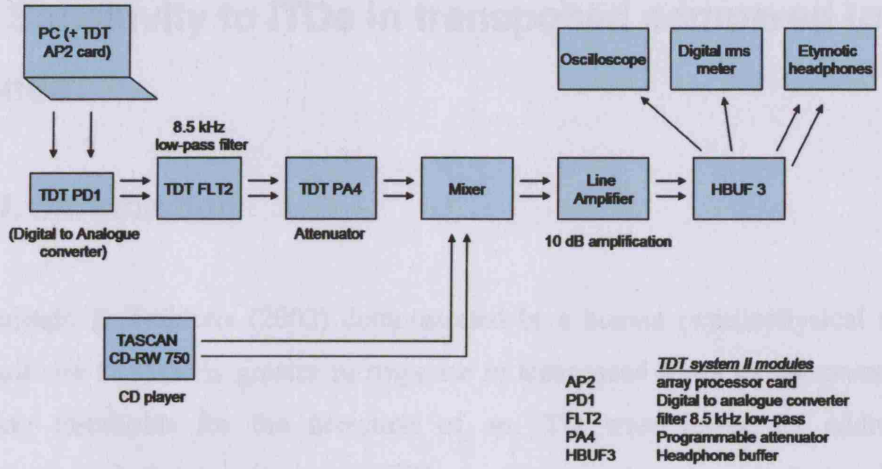

Figure 2.3. A diagram of the TDT system II hardware used for psychophysical experiments described in Chapter 5.

3. Sensitivity to ITDs in transposed compared to SAM tones

3.1. Introduction

Bernstein & Trahiotis (2002) demonstrated in a human psychophysical study that sensitivity to ITDs is greater in response to transposed tones as compared to SAM tones; thresholds for the detection of an ITD were lower. In addition, they demonstrated that sensitivity to ITDs in transposed tones can be equivalent to sensitivity to ITDs in low-frequency tones. In the experiments described in this Chapter, the hypothesis was tested that single neurons in the IC of the guinea pig show greater sensitivity to ITDs within transposed tones as compared to SAM tones. Neural thresholds for detection of an ITD in SAM and transposed tones were calculated and were compared to published neural thresholds previously calculated from the responses of low-CF neurons to pure tones.

3.2. Data Analysis

Data were obtained from 82 high-CF neurons from a total of 34 animals.

Sensitivity to ITDs in transposed and SAM tones

Sixty-nine neurons were investigated using SAM and transposed tones of 500-ms duration with ITDs generated by delaying the entire waveform. Spike rates were calculated by counting the number of spikes occurring over the middle 350 ms of the stimulus presentation, and over a time period that ensured a whole number of modulation periods within the 350 ms. This middle time period was employed in order to exclude neural responses to the onset of the stimuli, which had higher firing rates for some neurons, and to eliminate response to periods of monaural stimulation at large values of ITD. Periods of monaural stimulation could be up to 50 ms; for the

lowest modulation rate used (10 Hz) with an ITD equivalent to 0.5 cycles of interaural phase difference, the 500-ms stimulus was gated on at one ear 50 ms before the stimulus was gated on at the other ear, and gated off 50 ms before the stimulus offset. ITD-sensitivity was also assessed using spike rates calculated from a 600 ms analysis window which included all stimulus-driven spikes. This is described in Section I.1.1.

In a small number of early experiments the stimulus duration was 800-ms, or was 500-ms but did not have an onset-ITD. Spike rates from these experiments (data from 13 neurons) were calculated over the same 350 ms of the stimulus that was used for analysis of responses to 500-ms stimuli. The responses appeared qualitatively similar to those recorded from neurons using stimuli of 500 ms and are therefore included in the data presented in this Chapter.

Sensitivity to ITDs was examined by calculating the mean spike rates at each f_m and over ± 0.5 cycles of IPD. The following calculations were performed for each f_m (i.e. for each *recording*) in response to SAM and transposed tones. The mean phase vector was calculated using the method of Goldberg and Brown (1969). At every f_m, two measurements for an IPD of 0.5 cycles were taken (at +0.5 and -0.5 cycles) and the mean of these was used in the calculation. (See Section I.2 for a discussion of mean firing rates in response to -0.5 and 0.5 cycles, which were not significantly different.) The mean phase vector was termed the peak-IPD; this was considered the IPD to which the recording was tuned. The normalized length of the vector (the vector strength, which varied between 0 and 1) was used to calculate the Rayleigh coefficient (Kuwada *et al.*, 1987). The Rayleigh coefficient was the test static for a chi-squared distribution with two degrees of freedom (Mardia & Jupp, 1999). An individual recording was classified as "ITD-sensitive" if the Rayleigh coefficient was greater than 13.815, $p < 0.01$ (Kuwada *et al.* 1987; Mardia & Jupp, 1999), providing evidence that firing rates were not evenly distributed around a cycle of IPD. Individual *neurons* were classified as ITD-sensitive if one or more recordings were classified as ITD-sensitive.

High-resolution functions

High-resolution IPD functions were obtained in order to calculate neural discrimination thresholds. ROC (Receiver Operating Characteristic) analysis was used to estimate the smallest ITD that could be discriminated from 0-μs ITD, using the firing rates of a single neuron. IPDs were converted into ITDs with reference to the f_m. ROC analysis has been used previously to compare neural ITD-sensitivity to human psychophysical performance (Skottun *et al.*, 2001; Shackleton *et al.*, 2003). For this analysis, spike rates were calculated over the middle 350 ms analysis period. To perform ROC analysis, a neurometric function was calculated for each high-resolution IPD function. The distribution of firing rates obtained at 0-μs ITD was compared with the distribution of firing rates at every IPD presented. The spike rates recorded from all repetitions (6-10) of each ITD form a sample distribution of the spike rates at that ITD. If 101 ITDs were presented, 101 pairs of spike rates are compared, including the comparison of the spike rates at 0-μs ITD with itself. The neurometric functions describe the probability of randomly selecting a spike rate from each pair of distributions and finding the spike rate at 0-μs ITD to be lowest.

Neurometric functions were smoothed by averaging over three consecutive ITDs. ITDs were considered discriminable from 0-μs ITD if the probability (from the smoothed function) was >= 0.75 or <= 0.25 (giving two discrimination thresholds). The just noticeable difference (JND) in ITD was defined as the smallest ITD that is discriminable from 0-μs ITD. No interpolation was carried out to estimate the point where the neurometric functions crossed 0.75 or 0.25. JNDs were, therefore, if anything, slightly overestimated. For 5 high-resolution functions, 102 (instead of 101) values of IPD were presented. In these cases no response to 0-μs ITD was obtained, and discrimination was then calculated from the closest negative ITD to 0-μs.

Period histograms

Entrainment of spike times to the modulation period for SAM and transposed stimuli was examined by binning spike times over one cycle of the modulation rate. Spike times over the middle 350 ms of the stimulus presentation, and over a time period that ensured a whole number of modulation periods within the 350 ms, were analyzed. The cycles of the modulation period at which spikes occurred were adjusted such that 0 cycles was always with reference to the onset of sound in the contralateral ear. To quantify the degree of phase-locking to the stimulus period, at each IPD and f_m, the vector strength was calculated. Phase-locking was considered significant if the Rayleigh coefficient was > 13.815 ($p < 0.001$; Mardia & Jupp, 1999).

3.3. Results

3.3.1. Sensitivity to envelope ITDs in SAM and transposed tones

Responses to SAM and transposed tones were obtained from 82 IC neurons with CFs $\geq$ 2 kHz (2.0 – 13 kHz, median = 4.5 kHz). For 55/82 neurons, peri-stimulus time histograms (PSTHs) were obtained in response to monaural and diotic CF tones. All neurons exhibited greater firing rates in response to tones at the contralateral ear than at the ipsilateral ear. For 70% of the sample (39/55), ipsilateral stimulation resulted in inhibition, exhibited as either a reduction in the spontaneous firing rate in response to ipsilateral stimulation or as a reduction in the rate of firing in response to a contralateral sound when stimulation was binaural.

Responses of a typical IC neuron sensitive to ITDs in the envelope of SAM and transposed tones are shown in Figure 3.1. The neuronal CF was 3.4 kHz, which was confirmed by the frequency-vs.-level response area (Figure 3.1A). PSTHs for monaural and diotic CF tones indicate the neuron to be excited by contralateral stimulation but unresponsive to ipsilateral stimulation (Figure 3.1B). The response to

diotic stimulation was characterised by a peak spike rate lower than that to contralateral stimulation alone, but longer in duration.

Raster plots (Figures 3.1C-H) show responses to SAM or transposed tones which were interaurally delayed over a range of ITDs equivalent to interaural *phase* differences ± 0.5 cycles of the f_m, for f_m between 10 Hz and 510 Hz. Because the range of ITDs encompassed by ± 0.5 cycles of IPD differs for different modulation rates, all functions are plotted with respect to IPD. Phase-locking to the envelope of the stimulus was evident in the responses to both stimuli. The neuron was broadly tuned for IPD, favouring IPDs around 0 cycles. At 110 Hz f_m, IPD tuning was enhanced compared with 10 Hz, with discharge rates increasing at favourable IPDs and decreasing at unfavourable IPDs. The response at favourable IPDs was greater for transposed than for SAM tones. As f_m increased, discharge rates fell, particularly during the latter portion of the stimulus, and responses became less modulated with envelope delay. The criteria for ITD-sensitivity (see Section 3.2) were met in response to transposed tones at f_m between 10-310 Hz and for SAM tones at 110 and 210 Hz. Although the entire SAM and transposed waveforms were delayed, ITD sensitivity was conveyed by the envelope structure of the stimuli; neurons are not sensitive to carrier-based ITDs at frequencies above 2 kHz (Joris, 2003). The modulation in firing rates occurred over a cycle of interaural phase with respect to the modulation frequency and not the carrier frequency.

Responses to SAM (left) and transposed (right) tones across all f_m and IPDs are displayed as 3D mesh plots in Figure 3.1I. The neuron had a band-pass rate modulation transfer function (rMTF), with maximum firing rates at 110 Hz, a preference for IPD close to 0 cycles, and enhanced responses to transposed tones compared with SAM tones; peak discharge rates at the preferred f_m were almost double those evoked by SAM tones.

Responses to IPDs imposed on SAM and transposed stimuli for a second IC neuron, with a CF of 4.3 kHz, are shown in Figure 3.2. PSTHs in Figure 4B indicate that

contralateral stimulation evoked higher discharge rates than binaural stimulation; ipsilateral stimulation caused discharge rates to fall below the spontaneous rate of the neuron. Raster plots (Figure 3.2C-H) show spike times for six recordings in response to both SAM and transposed tones from 10 to 650 Hz f_m. Spike times were clearly phase-locked to the envelope at 10 Hz modulation. The neuron was sensitive to IPDs in both stimuli at 10 Hz; the pattern of the phase-locking to the modulation period changed with IPD and the spike rate fell around 0-cycles IPD. At 138 Hz f_m (Figure 3.2D), the neuron was IPD-sensitive to transposed, but not to SAM, tones. The neuron was insensitive to IPDs in either stimulus at the higher modulation rates (Figure 3.2E-H).

Figure 3.1I indicates that, at 10 Hz f_m, discharge rates were greater at favourable IPDs and lower at unfavourable IPDs, for transposed tones than for SAM tones. IPD sensitivity was low-pass with respect to f_m, unlike the band-pass response in Figure 3.1I. The responses also differ from those in Figure 3.1 in another important way. In Figure 3.2, discharge rates were minimal for values of IPD near 0 cycles whereas they were maximal near 0-cycles IPD in Figure 3.1. Specifically, the mean peak-IPD in response to transposed tones was 0.39 cycles at 138 Hz f_m (equivalent to an ITD of +2826 μs) whereas, in Figure 3.1, the mean peak-IPD was 0.03 cycles at 110 Hz f_m (equivalent to an ITD of +272 μs).

Examples of recordings from two ITD-*in*sensitive neurons are shown in Figures 3.3 and 3.4. The neuron referred to in Figure 3.3 (CF = 3.1 kHz) exhibited a strong onset response to diotic tones followed by a pause and then a more sustained response (Figure 3.3B). The PSTH to contralateral stimulation was similar but with a less well-defined pause, whereas ipsilateral stimulation resulted in a small onset response followed by inhibition of spontaneous firing (Figure 3.3B). Raster diagrams (Figures 3.3C-H) and the 3D mesh plot (Figure 3.3I) indicate that the neuron's discharge rate increased with increasing f_m. That is, discharge rate was high-pass with respect to f_m. At higher values of f_m, the discharge rate was greater for transposed than for SAM

tones while at the lower values of f_m, the opposite was the case. Thus, the discharge rate exhibited greater modulation as a function of f_m in response to transposed tones. Visual inspection suggests some form of binaural interaction occurred at f_ms of 10 and 20 Hz; phase-locking appeared weaker near 0-cycles IPD (Figures 3.3C & D). Further, at 320 Hz and below, the discharge rate was lower for values of IPD near 0 cycles but the criteria for ITD-sensitivity were not met (Figure 3.3I).

Neural responses in Figure 3.4 (CF = 2.8 kHz) were also insensitive to ITDs in SAM or transposed tones. PSTHs to pure tones (Figure 3.4B) indicate the neuron to be excited by contralateral stimulation and the number of spikes to be facilitated by binaural stimulation. The PSTH to ipsilateral stimulation shows inhibition of spontaneous firing, following a small onset response. In contrast to the onset response of the neuron in Figure 3 and the onset-pause-sustained responses in Figures 3.2B and 3.3B, this neuron had an adapting PSTH to binaural stimulation. Raster plots in Figure 3.4C-H indicate strong phase-locking to the f_m. At 10 Hz (Figure 3.4C) it was clear that the spike times were more tightly phase-locked to the f_m in response to transposed tones than SAM tones. Although discharges were entrained to the envelopes of both SAM and transposed tones, the response to the transposed tones occurred over a shorter duration of the modulation period, reflecting the shorter "on-period" (where the value of the sound pressure waveform is > 0) per cycle of the transposed tone. Mean firing rates in Figure 3.4I were not modulated with either f_m or IPD.

Mean firing rates of a further nine neurons, chosen to illustrate the variety of responses obtained, are displayed as 3D mesh plots in Figure 3.5. Neurons classified as sensitive to envelope-based ITDs and with firing rates that peak close to 0-cycles IPD are arranged in the top row and neurons with firing rates that peak close to IPDs of 0.5 cycles in the middle row. The bottom row shows neurons insensitive to envelope-based ITDs. For each ITD-sensitive neuron, the maximum Rayleigh coefficient (which was used as a measure of sensitivity to ITDs) was determined across all recordings in response to SAM and transposed tones and both the top and

middle rows (Figures 3.5A-C and D-F) are arranged from left to right in order of decreasing maximum Rayleigh coefficient.

All three neurons in the top row of Figure 3.5 showed greater modulation of their discharge rate in response to transposed tones than SAM tones, as a function of both IPD and f_m. The neuron in Figure 3.5A was very strongly sensitive to ITDs conveyed by transposed tones over a specific range of modulation rates (138 to 266 Hz), with the greatest Rayleigh coefficient at 266 Hz, whereas it was largely unresponsive to ITDs conveyed by the SAM stimulus. The other two neurons (Figures 3.5B & C) responded strongly to both transposed and SAM tones and were sensitive to ITDs over a specific range of f_m (around 20 & 40 Hz in Figure 3.5B and around 115 & 220 Hz in 3.5C in response to transposed tones and around 40 Hz in Figure 3.5B in response to SAM tones). Generally, when neurons were sensitive to ITDs in both stimuli, sensitivity occurred at the same f_m but extended to a wider range of f_m in response to transposed tones (e.g. in 3.5B). Overall, discharge rates were more highly modulated to ITDs conveyed by transposed stimuli than they were to ITDs conveyed by SAM stimuli. In addition, rates were consistently higher at favourable ITDs and lower at unfavourable ITDs for ITDs conveyed by transposed as compared to SAM tones. Tuning to IPDs, at f_m to which sensitivity to ITDs occurred, was similar in response to both stimuli i.e. peaks and troughs in firing rates occurred at similar IPDs in response to SAM and transposed tones. As a function of f_m, the neural response in Figure 3.5C had a trough in its firing rate at 115 Hz f_m which was more pronounced (i.e. reaching a lower firing rate) in response to transposed tones than SAM tones.

Discharge rates as a function of IPD were also more modulated in response to transposed tones for the responses in Figures 3.5D & E. Neurons were insensitive to ITDs in the envelope of SAM tones whereas in Figures 3.5D & E, sensitivity occurred for more than one f_m in response to transposed tones. Atypically in the sample of neurons which were recorded from, the firing rate of the neuron in Figure 3.5F was more modulated as a function of IPD in SAM tones as compared to transposed tones at 250 Hz f_m (although the reverse was true at 330 Hz). There was a

response minimum at 250 Hz f_m which was lower for transposed tones and may explain the reduced modulation in firing rate as a function of IPD to transposed tones. As in Figures 3.5B & C, the IPD tuning in Figures 3.5F was similar for SAM and transposed tones, i.e. peaks and troughs in firing rate were similarly positioned and, for each neuron, there was a modulation rate at which sensitivity was greatest (210 Hz in 3.5D, 10 Hz in 3.5E and 250 Hz in 3.5F).

Raleigh coefficients were higher in response to transposed tones as compared to SAM tones (median = 3.3 vs. 1.5, Wilcoxon paired signed-rank test, $z = 11.7$, $p < 0.001$). More recordings were classified as sensitive to ITDs within transposed tones than in SAM tones (Table 1; McNemar's test, χ^2 (1 d. f.) = 110.25, $p < 0.001$). Correspondingly, more neurons were sensitive to ITDs within transposed than SAM tones (Table 1; McNemar's test, χ^2 (1 d. f.) = 18.375, $p < 0.001$). Neurons were broadly classified as V- (46%), I- (47%) or O- (7%) type according to any inhibitory side-bands in their frequency vs. level response areas (Ramachandran *et al.* 1999). This classification was not predictive neurons' sensitivity to SAM or transposed tones.

Neurons for which responses are shown in the bottom row of Figure 3.5 were not sensitive to envelope-ITDs in SAM or transposed tones. However, their discharge rates were modulated as a function of f_m. For each of the three neurons, changes in discharge rate as a function of f_m were more pronounced in response to transposed tones than SAM tones. Although not of central interest to the current study, neural sensitivity to the f_m was examined, in the form of rMTFs (see Section I.1.4). Over all cells, rMTFs were more modulated in response to transposed as compared to SAM tones.

3.3.2. Sensitivity to ITD as a function of f_m

The number of ITD-sensitive recordings as a proportion of the total number of recordings is displayed in Figure 3.6A, calculated as a function of f_m, in 50-Hz bins. Sensitivity to ITDs could be described as band-pass as a function of f_m, with the greatest proportion of ITD-sensitive recordings between 60 and 310 Hz f_m. The sharpest reduction in the proportion of ITD-sensitive recordings occurred as the f_m increased above 310 Hz and no recordings were considered sensitive to ITDs at f_m above 550 Hz. The total number of recordings in each f_m bin is shown in the inset of Figure 3.6A. The f_ms at which SAM and transposed tones were presented were limited to $\leq$ CF/5. However, there was no relationship between the highest *fm* at which sensitivity to ITDs occurred and the CF of the neuron (Figure 3.6B). For each neuron, the f_m at which sensitivity to ITDs was greatest was estimated by calculating the synchronization rate (vector strength x mean firing rate) of each ITD-sensitive recording (Yin *et al.*, 1986). The distribution of the f_m at which sensitivity to ITDs was greatest is shown in Figure 3.6C. Most of the f_ms at which the synchronization rate was highest were below 210 Hz.

The distribution of peak-IPDs for all ITD-sensitive recordings is shown in Figure 3.7. Positive peak-IPDs indicate a preference for sounds leading at the contralateral ear; negative peak-IPDs indicate a preference for sounds leading at the ipsilateral ear. Peak-IPDs in response to both transposed tones (right) and SAM tones (left) clustered around 0 and 0.5 cycles, with more positive than negative values. This is consistent with previous physiological investigations of neural sensitivity to ITDs which indicate that IC neurons have, predominantly, peak-IPDs corresponding to sounds leading at the contralateral ear (Yin & Kuwada, 1983b; Yin *et al.*, 1984; Yin *et al.*, 1986; Caird & Klinke, 1987; McAlpine *et al.*, 2001).

For some ITD-sensitive neurons in this study it was possible to make recordings that were closely spaced in f_m, and the ITD-tuning of these neurons as a function of f_m was

investigated. Responses from the eight neurons with the highest number of ITD-sensitive recordings in response to transposed tones are shown in Figure 3.8A. Discharge rates were normalized to maximum at each f_m to emphasize the ITD-tuning. Responses to SAM tones are shown on the left and responses to transposed tones are shown on the right. Note that abscissae in Figures 3.8Ad&f are scaled differently compared to the other Figures in 3.8A to take account of the range of f_m at which sensitivity to ITDs was observed. Neurons are organized by eye, according to whether maximum discharge rates were aligned around a common ITD for different f_ms (top two rows) or whether minimum discharge rates were aligned around a common value of ITD for different f_ms (bottom row). For most neurons, the positions of response maxima and response minima as a function of f_m were similar for SAM and transposed tones.

Neurons in which response maxima are aligned are referred to as "peak-type" neurons and neurons in which response minima are aligned are referred to as "trough-type" neurons (Yin & Kuwada, 1983b; Batra *et al.*, 1997b). Phase plots (peak-IPD plotted as a function of f_m) are shown in Figure 3.8B for the same neurons for which responses are shown in Figure 3.8A. Theoretically, for neurons described as "peak-type" or "trough-type" neurons, Peak-IPD changes systematically with pure-tone frequency or f_m (Rose *et al.*, 1966; Joris, 1996; Batra *et al.*, 1997b). The slope of a linear regression describing the relation between Peak-IPD and frequency is referred to as the characteristic delay (CD) and the intercept with the ordinate axis is referred to as the characteristic phase (CP; Rose *et al.*, 1966; Yin & Kuwada, 1983b). For "peak-type" neurons, the CD is equal to the ITD at which the neuron responds maximally. A CP of 0 would be predicted from perfect coincidence of excitatory (EE) inputs. A CP of 0.5 would be predicted from coincidence of excitatory and inhibitory inputs (EI). Although it was possible to fit a linear regression to some of the phase plots in Figure 10B (e.g., Figures 3.8Bd & h in response to SAM tones, and in Figure 3.8Ba in response to transposed tones have $R^2 > 0.6$), for other neurons the linear regression was not appropriate (e.g., Figure 3.8Bd). Previous investigations of sensitivity to ITDs of pure tones and high-frequency SAM tones in the IC reported a

similar range of neural responses; phase plots are often not well-fitted by linear regression and, where phase plots are well fitted by linear regression, CPs may not be close to 0 or 0.5 (Yin & Kuwada, 1983b; Kuwada *et al.*, 1987; Batra *et al.*, 1993; McAlpine *et al.*, 1996; Batra *et al.*, 1997b).

3.3.3. Discrimination thresholds for ITDs in the envelope of transposed tones: comparison with SAM tones and low-frequency pure tones.

Of primary interest to this investigation was the neural sensitivity to ITDs in SAM as compared to transposed tones. In order to quantify the sensitivity to ITDs in response to transposed tones and SAM tones, high-resolution IPD functions were obtained and neural discrimination thresholds were calculated using ROC analysis (see Section 3.2). Neural discrimination thresholds of ITDs in low-frequency tones have previously been calculated by Shackleton and colleagues (Skottun *et al.*, 2001; Shackleton *et al.*, 2003). The lowest thresholds obtained were comparable to human performance in ITD detection tasks (Shackleton & Palmer, 2004).

High-resolution IPD functions were obtained in response to SAM and transposed tones from 14 neurons. For 12 neurons, functions were obtained only at a single f_m. For two neurons, functions were obtained at two f_ms (for one of these two neurons, a second high-resolution function was obtained in response to transposed, but not SAM, tones). This gave a total of 16 functions obtained in response to transposed tones, and 15 functions obtained in response to SAM tones. Six examples of such functions are shown in Figures 3.9A-F. On the left, the mean discharge rates and standard deviations (error bars) are shown in response to SAM (blue) and transposed (red) tones. The abscissae differ across plots because they reflect the presentation of different values of f_m which, given the constant range of IPDs presented, resulted in different ranges of ITDs being presented. Neurons in Figures 3.9A, C & E responded best to IPDs around 0 cycles whereas those in Figures 3.9B, D & F responded best to

IPDs around 0.5 cycles. Consistent with the responses shown in Figures 3.1, 3.2 & 3.5, discharge rates were more highly modulated to transposed tones as compared to SAM tones.

The probability of correctly discriminating between 0-□s ITD and all other ITDs presented is shown in the neurometric functions in each of Figures 3.9A-F (right panels). Thresholds for discrimination (see Section 3.2) are marked by black horizontal lines. From these ITDs, the single value closest to 0 □s was taken to be a measure of the Just Noticeable Difference (JND). Neural JNDs are shown by dashed lines for SAM (blue) and transposed (red) tones.

Figure 3.9A depicts data obtained with a f_m, of 100 Hz. At that rate of modulation, the neural JNDs were 644 μs for SAM, and 258 μs for transposed tones. Figure 3.9C shows data obtained from the same neuron in Figure 3.9A, but for a f_m of 140 Hz. For that rate of modulation, JNDs were 672 μs and 177 μs for SAM and transposed tones, respectively. Consistent with the greater modulation in firing rates in response to transposed tones as a function of ITD, JNDs derived from these functions were correspondingly lower for transposed than for SAM tones. This was also the case in Figures 3.9B, D and E. For the responses shown in Figure 3.9E, the neural JNDs were large, being 10 ms and 3.3 ms in response to SAM and transposed tones, respectively. These large values were, at least in part, a result of the large variability observed in the measures of firing rate compared to the mean firing rate change over a cycle of IPD. For the responses depicted in Figure 3.9F, the JND in response to transposed tones was > 2 ms and, in response to SAM tones there were no ITDs that were determined to be discriminable from an ITD of 0 μs.

JNDs ≤ 4 ms for SAM and transposed tones are plotted as a function of f_m in Figure 3.9G. All 16 JNDs calculated in response to transposed tones were ≤ 4 ms and are marked by red diamonds in Figure 3.9G. For two neurons, JNDs were calculated from high-resolution functions at two f_ms. These are indicated by a square or a

circular black outline around the red diamonds. For one of these neurons (the JND marked by a circular outline) a high-resolution IPD function was obtained in response to SAM tones at only one f_m. Of the 15 JNDs obtained in response to SAM tones, only 10 were $\leq$ 4 ms, and these are plotted in Figure 3.9G as blue circles. The remaining 5 JNDs were either greater than 4 ms or were immeasurable. The horizontal dotted lines in Figure 3.9G indicate estimates of the maximum ITD experienced by the guinea pig. The lower value is from a theoretical consideration of the maximum difference in time for sound travelling around a spherical head (McAlpine *et al.*, 2001) and the higher value is from head-related transfer functions (HRTFs) measured in the guinea pig (Sterbing *et al.*, 2003; see Section 3.4).

Also shown in Figure 3.9 are neural JNDs for ITD measured with low-frequency tones below 400 Hz (Shackleton *et al.*, 2003; Shackleton & Palmer, 2004) for low-CF neurons in the guinea-pig IC (squares). Median JNDs in response to SAM (blue), transposed (red) and low-frequency tones (black), are marked by asterisks. Overall, JNDs in ITD are lower for transposed tones than for SAM tones (Mann-Whitney U test for equal medians, $p < 0.0087$). In the case of just one neuron (also shown in Figure 3.5F), and only at a f_m of 200 Hz, was the JND lower in response to SAM tones than transposed tones (0.5 ms vs. 1.4 ms). When f_m was increased to 280 Hz, the JND for that same neuron was 320 μs in response to transposed tones, lower than either JND at 200 Hz.

The lowest JNDs in response to transposed tones were comparable to JNDs obtained in response to low-frequency tones, at corresponding frequencies, whereas the lowest JNDs in response to SAM tones were substantially higher. For f_m over 280 Hz, no neural JNDs could be calculated. As previously discussed (see Figure 3.6), responses indicating sensitivity to ITDs were rarely recorded at these rates and when such responses were observed they were only weakly modulated with variations in ITD. In contrast, neural JNDs in response to low-frequency tones *decrease* as the tone frequency increases up to ~500 Hz (Shackleton & Palmer, 2004). JNDs in response to

low-frequency tones between 280 and 400 Hz contribute to the low median JND for low-frequency tones.

3.3.4. Phase-locking to the envelope modulation

Transposed tones are designed to evoke a similar temporal pattern of action potentials from high-frequency ANFs to that evoked in low-frequency fibres by low-frequency pure tones. As previously discussed, high-frequency ANFs would not be expected to show this pattern in response to SAM tones. It is predicted, therefore, that ANFs should show "tighter" phase-locking to the period of a transposed than a SAM waveform (compare the right-hand-side of Figures 1.7.1B and 1.7.1C). Phase-locking to the f_m was examined in the responses of the sample of IC neurons recorded from.

Period histograms of the number of spikes occurring throughout the period of SAM and transposed waveforms were calculated for the four example neurons in Figures 3.1 – 4 and are displayed in Figure 3.10A-D, respectively. Paired responses to SAM (top) and transposed (bottom) tones are shown in response to each modulation rate, with the response to 10 Hz modulation on the left. The response was clearly phase-locked to 10 Hz in Figure 3.10A. In response to transposed tones, spikes occur over a narrower range of the stimulus cycle than they do in response to SAM tones. In the second pair of histograms, measured at 110-Hz modulation, phase-locking occurred in response to both stimuli but was tighter in response to transposed tones. As the modulation rate increased, phase-locking became weaker and was no longer apparent at 310 Hz. A similar pattern was observed in Figures 3.10B, C & D; the phase-locking was limited to the lower modulation rates in response to both stimuli and in response to transposed tones was "tighter" than to SAM tones. This likely reflects the shorter "on-period" per cycle of the transposed tones.

To summarize phase-locking to the modulation period for all neurons in our sample, the vector strength was calculated at each value of f_m and IPD that yielded

significantly phase-locked responses (Rayleigh coefficient > 13.815, $p < 0.001$). A three factor ANOVA for stimulus type, f_m and IPD revealed each factor had a significant effect on the phase-locking to the modulation rate (all $p < 0.001$.). Changes in phase-locking, as a function of IPD, can be seen in the raster plot in Figure 3.2C.

In order to examine the effect of stimulus type and f_m on the vector strength, the vector strength values (for all recordings that showed significant phase-locking) were grouped into logarithmically-spaced f_m bins. The means for each f_m bin are shown in Figure 3.11. Overall, phase-locking was greater in response to transposed than SAM tones (t-test, $p < 0.001$) and was low-pass as a function of f_m for both stimuli. The vector strengths in response to both stimuli decreased from similar f_ms (around 100 Hz) and the characteristics of the reduction were qualitatively similar. Phase-locking is essential for coding temporal information that mediates sensitivity to ITDs in binaural neurons. The reduction in phase-locking with increasing f_m might underlie the limit on sensitivity to envelope-based ITDs observed at around 300 Hz (see Figure 3.6 & 3.8 and Section 3.4).

3.4. Discussion

The principal finding from the experiments presented in this Chapter is that neural sensitivity to ITDs conveyed by the envelopes of high-frequency sounds can be as great as that observed for ITDs conveyed by the fine-structure of low-frequency sounds. The neural data presented here, recorded from the IC of the guinea pig, are consistent with the enhanced ITD-sensitivity yielded by transposed tones observed in psychophysical data with human listeners (Bernstein & Trahiotis, 2002). A greater number of neurons were sensitive to ITDs within the envelope of transposed as compared to SAM tones. Firing rates were more modulated as a function of ITD and a discrimination analysis demonstrated that JNDs for ITD were lower in response to

transposed tones as compared to SAM tones. It was concluded that neurons are more sensitive to ITDs within transposed tones than SAM tones.

3.4.1. Comparison with human psychophysics using transposed tones

The experiments presented in this Chapter were motivated by the finding that, for human listeners, transposed tones confer an improvement in sensitivity to ITDs at high carrier frequencies compared to SAM tones (Bernstein & Trahiotis, 2002). The neural JNDs followed a similar pattern to those obtained from human listeners; threshold-ITDs obtained with transposed tones were smaller than those obtained with SAM tones. It is noteworthy that Bernstein and Trahiotis (2002) found threshold-ITDs to be immeasurable with the high-frequency stimuli for rates of modulation above 256 Hz f_m. A similar limitation was observed in the neural data presented in the current study.

Clearly, one should be cautious when comparing human psychophysical data with neural recordings from guinea pigs. Human thresholds for detection of ITDs were obtained during a behavioural task in which the subjects were awake and were motivated to undertake the task assigned. It is not known which (or how many) neurons in the human brain were needed to accomplish this task. In contrast, neural thresholds were obtained from single neurons in the inferior colliculus of the anaesthetised guinea pig. Differences in thresholds may therefore be expected due to differences between species or differences in attentional state. The neural JNDs are a measure of how much information there is in the response of a single neuron. It is not suggested that single neurons are necessarily sufficient to carry out detection of ITDs. Behavioural thresholds may require pooling of neural responses. It is interesting that the JNDs obtained from the most sensitive neurons follow a similar pattern, across frequency and stimulus type, to human psychophysical thresholds. It is suggested that the ability to code ITDs of single neurons in the human IC would also be related to

human psychophysical performance. Conversely, if guinea pig behavioural experiments were carried out, it is predicted that similar limitations in detections of ITDs would be shown as are indicated by the neural JNDs and the human psychophysical thresholds.

3.4.2. Comparison with previous *in vivo* electrophysiology

There are currently no other published reports of ITD-tuning functions in response to transposed tones. Previous authors have reported mammalian neural sensitivity to ITDs in the envelopes of high-frequency stimuli including data using SAM tones recorded from the MSO (Yin & Chan, 1990; Batra *et al.*, 1997a; Batra *et al.*, 1997b), LSO (Joris & Yin, 1995; Joris, 1996) and IC (Batra *et al.*, 1989; Batra *et al.*, 1993), noise recorded from the IC (Joris, 2003), clicks recorded from the IC (Caird & Klinke, 1987) and high-frequency tones with trapezoid envelopes recorded from the IC (Yin *et al.*, 1984). The neural responses to SAM and transposed tones, in the current study, were qualitatively similar to these previous studies; peak-IPDs were predominantly leading at the contralateral ear and phase plots were not necessarily well-described by linear regression. Clustering of peak-IPDs close to 0 and 0.5 cycles was reminiscent of EE and EI interactions (see Section II.3.7) described previously in response to high- (Batra *et al.*, 1997a) and low-frequency sounds (Yin & Kuwada, 1983b).

Neural JNDs obtained in response to SAM and transposed tones were compared with neural JNDs obtained in response to ITDs in low-frequency tones in the IC of the guinea pig by Shackleton and colleagues (Skottun *et al.*, 2001; Shackleton *et al.*, 2003; Shackleton & Palmer, 2004). The current study used a longer stimulus duration (500 ms vs. 50 ms) and a smaller number of repetitions (6-13 vs. 100) than were used by Shackleton and colleagues. Onset and offset responses were excluded in the calculation of spike rates for ROC analysis in the current study, in contrast to the inclusion of the whole stimulus-driven response in the calculation of low-frequency

JNDs by Shackleton and colleagues. The stimulus used by Shackleton and colleagues had no onset-ITD. Despite the differences in stimulus characteristics, neural JNDs in response to transposed tones were comparable to JNDs in response to low-frequency tones at f_m < ~300 Hz. This similarity between thresholds in response to pure tones and transposed tone below 300 Hz is consistent with human psychophysics (Bernstein & Trahiotis, 2002). Finally, JNDs calculated over the 600-ms analysis window, and the 350-ms analysis window, were similar. The effect of calculating JNDs from firing rates calculated over a 600-ms analysis window compared to a 350-ms analysis window is explored in Section I.1.5.

3.4.3. Phase locking and a f_m limitation on sensitivity to ITDs

It is suggested that the enhanced neural ITD sensitivity observed with transposed stimuli was a result of the temporal pattern of action potential firing in high-frequency ANFs, stemming from the temporal signature of the envelopes of those stimuli. Transposed tones were originally designed to overcome the nature of peripheral auditory processing *hypothesised* to be responsible for the relatively poor binaural performance observed with high-frequency stimuli (van de Par & Kohlrausch, 1997). The transposed stimulus was designed to produce an output at the level of high-frequency ANFs which, in terms of its temporal pattern, mimics the output normally observed at low-frequencies (Figure 1.7.1). The analysis of phase-locking in the IC presented here indicates that phase-locking was enhanced in response to transposed as compared to SAM tones (Figure 3.11). This enhancement in phase-locking does not require any increased ability to entrain to the transposed waveform but can be explained simply by the different envelopes of the stimuli. If a neuron's pattern of firing followed exactly the time course of SAM and transposed tones, the phase-locking to the transposed waveform would be "tighter" because the transposed waveform itself occurs over a narrow range of the modulation period.

The tighter phase-locking in response to transposed tones may explain the enhancement in sensitivity to ITDs observed in behavioural experiments with human listeners and in the physiological recordings presented here. Specifically, binaural neurons in the superior olivary complex (SOC) may receive temporally more precise patterns of action potentials in response to transposed tones than in response to SAM tones. Inputs arriving in phase at the SOC are more likely to produce coincident action potentials, with a high output from EE (peak-type) neurons, whereas inputs arriving out of phase produce fewer coincident action potentials, and reduced output from EE neurons. Responses to SAM tones were still highly phase-locked to the SAM waveforms and an alternative explanation for the improvement in ITD-sensitivity might involve, in some direct fashion, the "off period" in response to transposed and low-frequency tones. Binaural inputs might not be any more likely to arrive coincidently in response to transposed tones but rather the "off period" might be crucial for the efficient binaural processing of ITDs. In any case, the notion that the enhancement observed with transposed tones is directly tied to the properties of the stimulus is consistent with Bernstein and Trahiotis' (2002) finding that the enhancements observed psychophysically were well accounted for by changes in the normalized interaural correlation of the waveforms as processed by the auditory periphery.

Although transposed tones are designed to provide a pattern of responses in high-frequency ANFs similar to that seen in low-frequency fibres in response to low-frequency tones, the response patterns are, in fact, unlikely to be identical. In response to low-frequency tonal stimulation auditory hair cells are hyperpolarised during the "off" or non-preferred half-cycle of the motion of the basilar membrane, stereocilia are deflected towards their shortest row and the mechanically-gated ion channels are closed (Fettiplace & Fuchs, 1999). During these "off" periods, the response of an ANF can fall below its spontaneous rate. Although such a reduction would not be expected in response to high-frequency transposed tones, neural JNDs for modulation rates below 300 Hz are remarkably similar for transposed and low-frequency tones.

Data describing ANF responses to transposed tones have recently been published (Dreyer *et al.*, 2006). Phase-locking was found to be higher in response to transposed as compared to SAM measured by the vector strength and by the height of a shuffled autocorrelogram (Louage *et al.*, 2004). However, phase-locking was found to be strongly intensity dependent, decreasing as the stimulus intensity increased. Recordings were not regularly obtained in response to SAM and transposed tones at different intensities in the data presented in this thesis.

While phase-locking was greater in response to transposed tones, phase-locking decreases with f_m in the same manner in response to either SAM or transposed tones (Figure 3.11). This is consistent with the rate-limitation of phase-locking to amplitude modulated sounds measured previously (Joris & Yin, 1998; Krishna & Semple, 2000), and with behavioural accounts of *monaural* temporal modulation transfer functions (Ewert & Dau, 2000; Kohlrausch *et al.*, 2000). Not surprisingly given that ITD-processing is dependent on phase-locking, few neurons were ITD-sensitive to rates of modulation above 300 Hz (Fig 3.6 & 3.9). In contrast, ITD sensitivity in response to low-frequency tones occurs at frequencies above 300 Hz; human JNDs are minimal around 1 kHz (Klumpp & Eady, 1956) and guinea pig IC neural JNDs are minimal around 500 Hz (Shackleton & Palmer, 2004). Previous neural (Yin *et al.*, 1984; Batra *et al.*, 1997a; Joris & Yin, 1998) and psychophysical (McFadden & Pasanen, 1976; Nuetzel & Hafter, 1981; Bernstein & Trahiotis, 1994; Bernstein & Trahiotis, 2002) investigations at high-frequencies have observed a f_m limitation on sensitivity to ITDs.

As the modulation rate of an amplitude modulated stimulus increases such that the spectral components (see Figure 1.7.2) fall outside the spectral receptive field of an ANF, phase-locking to the amplitude modulation is lost. This forms an absolute limit on the modulation rates at which sensitivity to envelope ITDs is possible. For transposed tones, the middle 5 spectral components are required to gain a binaural advantage, psychophysically (van de Par & Kohlrausch, 1997). As f_m is increased,

therefore, the components that are most spectrally distant from CF are attenuated most. Thus, phase-locking and ITD sensitivity would be predicted to decrease when attenuation of spectral components occurs. The f_m at which ITD-sensitivity can occur would be predicted to increase with increasing CF because ANF spectral receptive fields widen with increasing CF (Evans *et al.*, 1992). Bernstein & Trahiotis (1994) tested this hypothesis in humans and found that this was not the case. Consistent with the findings of these psychophysical studies, the data from the present study revealed no relationship between CF and the maximum f_m at which ITD-sensitive recordings could be obtained (see Figure 3.6B). It therefore suggested that a neuron's spectral receptive field does not account for the observed limitation on the f_m at which sensitivity to ITDs in the envelope of high-frequency sounds can occur.

Joris and Yin (1998) proposed that a neural limit on sensitivity to amplitude modulation and to ITDs in high-frequency sounds arises in the LSO. It is likely that any such limitation results from the membrane properties of LSO neurons e.g. the presence or absence of ion channels capable of altering the ability of neurons to follow the temporal structure of their inputs (Barnes-Davies *et al.*, 2004).

3.4.4. Neural JNDs and physiological detection of ITDs

Natural sounds, such as human speech, have an envelope structure replete with onsets, offsets and "off-periods" (Rosen, 1992). ITDs conveyed by the envelope of high-frequency stimuli could be relevant to binaural tasks such as sound localisation and signal detection in noise (Best *et al.*, 2004). Would the neural JNDs measured in the current study be useful to the guinea pig? In Figure 3.9G, two measures of the "maximum" ITD that a guinea pig might experience are marked by dotted lines. The value of 330 μs is the maximum ITD, at low-frequencies, in HTRFs measured using noise stimuli by Sterbing *et al.* 2003 whereas the value of 150 μs is the result of a theoretical consideration of the time sound takes to travel around a spherical head (McAlpine *et al.*, 2001). Below 300 Hz, there was only one JND for low-frequency

tones that was <= 150 μs, and below 250 Hz there are few JNDs <= 330 μs. Similarly, in response to transposed tones, there are few JNDs below 330 μs. If single neurons are sufficient to encode ITDs, this suggests that the usefulness of these ITD-tuning functions is limited because the JNDs are close to the maximum ITD that can be experienced. Shackleton *et al.* (2003) found that the lowest JNDs in response to low-frequency tones were not generated using 0-μs ITD as a reference. In general, they found that JNDs were minimal with reference ITDs ipsilateral to the IC from which the neuron was recorded. Similarly, there may be regions of the IPD-tuning functions in response to transposed tones that have greater accuracy for ITD discrimination. It is not proposed that single neurons are sufficient to encode ITDs; it is more likely that the responses of many neurons are pooled.

Despite limitations in the temporal coding of neurons to high-frequency sounds, the results presented here demonstrate that under conditions in which binaural neurons receive appropriate spike patterns, sensitivity to ITDs conveyed by high-frequency stimuli can be equivalent to that observed in response to low-frequency stimuli. This suggests, as first conjectured by Colburn and Esquissaud (1976), that mechanisms underlying ITD sensitivity in low- and high-frequency channels of the auditory system are, to a first approximation, equivalent.

3.5. Figures and Tables for Chapter 3

	Total	Sensitive to ITDs in trans. Tones only	sensitive to ITDs in trans. and SAM tones	sensitive to ITDs in SAM Tones only	insensitive to ITDs
Number of neurons	82	23	23	1	35
Number of recordings	959	172	57	24	706

Table 1. The number of neurons and recordings sensitive to ITDs in SAM and transposed tones.

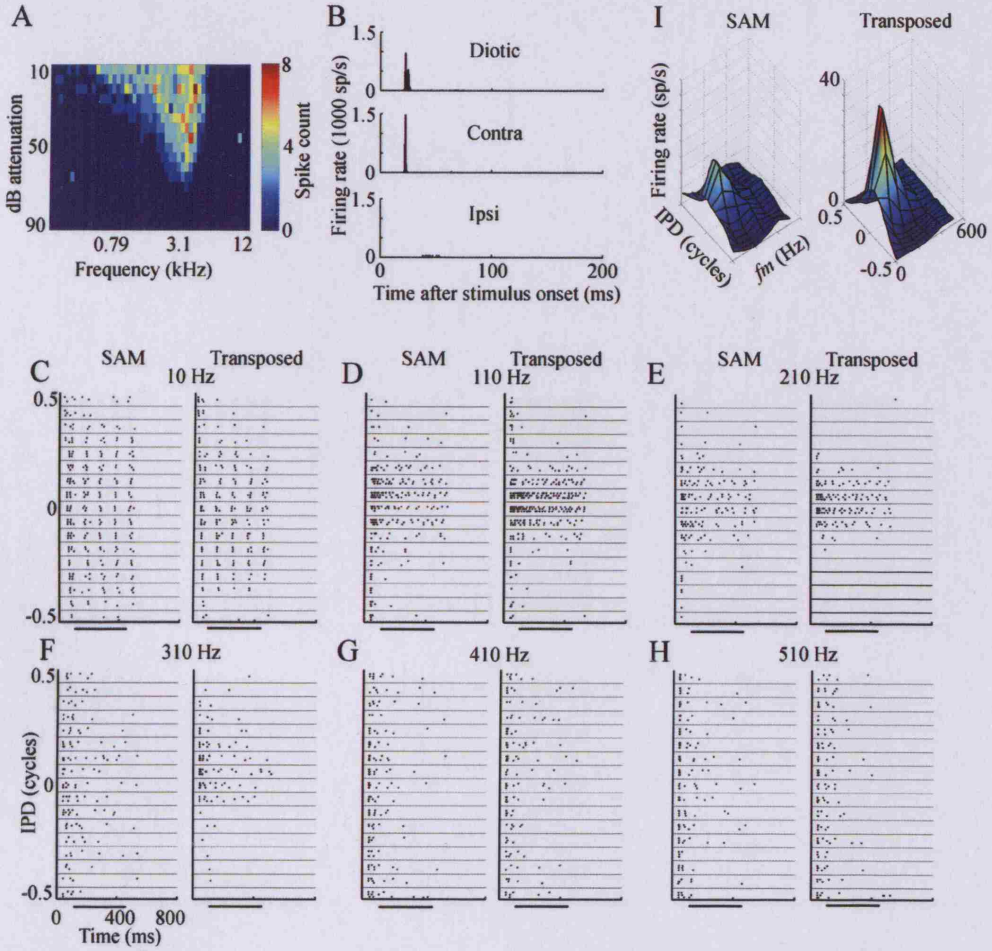

Figure 3.1. The responses of an IC neuron (28104; CF = 3.4 kHz; threshold = -67 dB re maximum system output) that was classified as ITD-sensitive to both SAM and transposed tones. (A) The neuron's frequency-versus-level response area. Spike counts are indicated by gray level. (B) PSTHs to 50 ms diotic, contralateral (Contra) and ipsilateral (Ipsi) tones at CF. (C) – (H) Raster plots for recordings at 10, 110, 210, 310, 410 and 510 Hz f_m respectively. All axes are scaled as in G. Each band on the ordinate shows spike times in response to repeat presentations of different IPDs from -0.5 to +0.5 cycles. *Black bars below the abscissae* indicate the middle 350 ms of the stimulus. (I) 3D mesh plots of mean spike rate, calculated over the period indicated by black bars in (C)-(H), as a function of IPD and f_m. Mean spike rates were averaged over 3 consecutive IPDs.

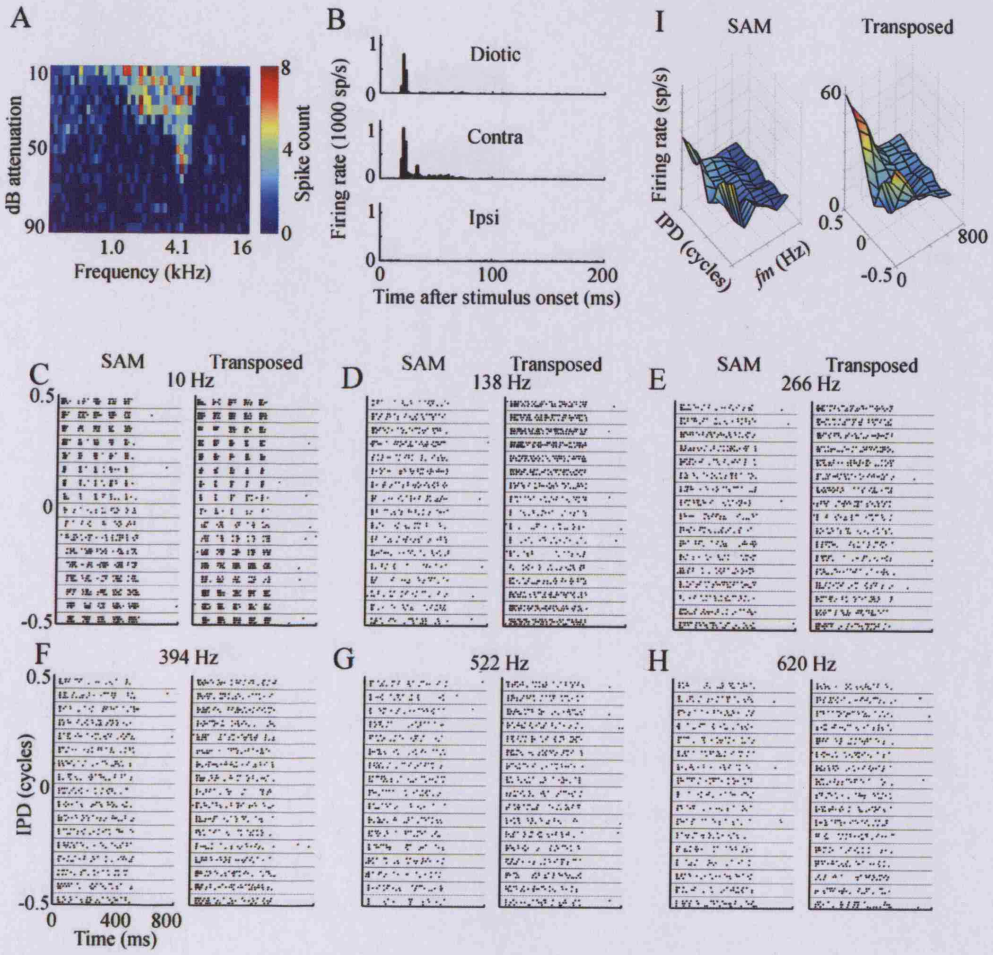

Figure 3.2. The responses of an IC neuron (25916; CF = 4.3 kHz; threshold = -68 dB re maximum system output) that was classified as ITD-sensitive to both SAM and transposed tones. The figure follows the same format as Figure 3.1. (C) - (H) Raster plots for recordings at 10, 138, 266, 394, 522 and 650 Hz f_m. Black bars indicating the middle 350 ms of the stimulus are not shown.

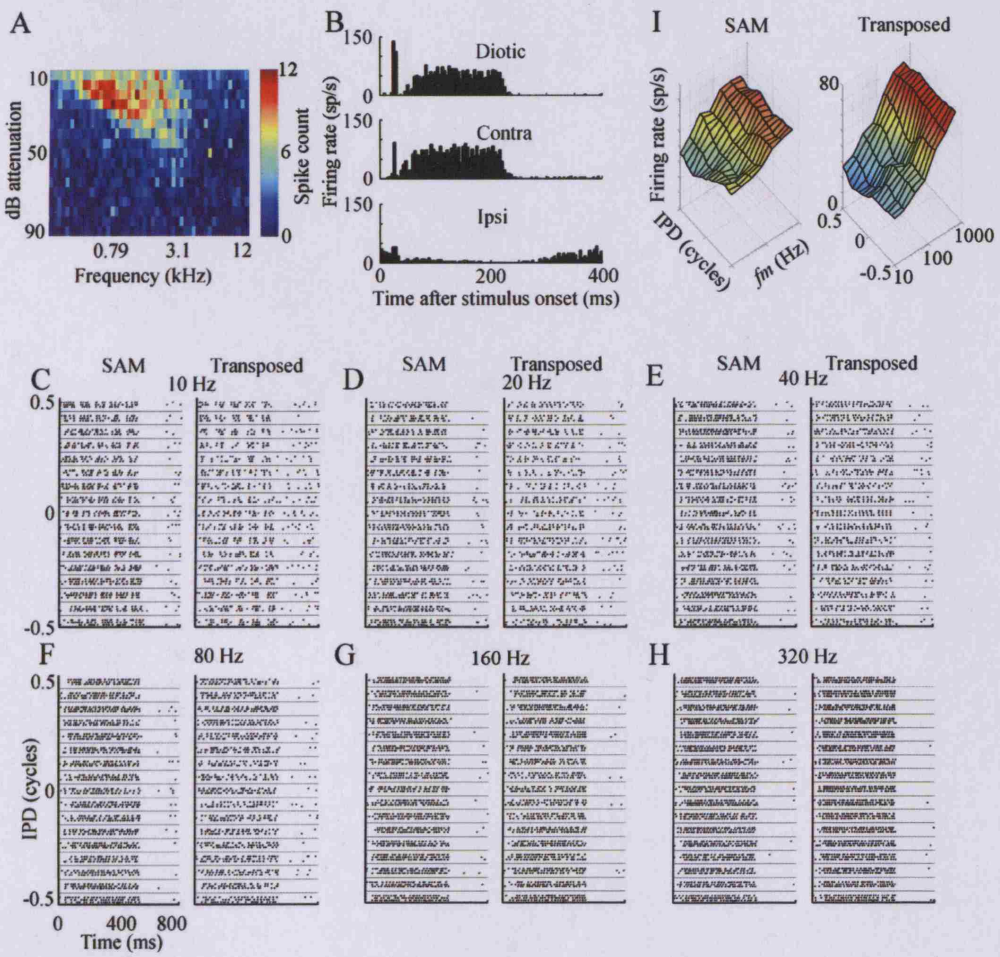

Figure 3.3. The responses of an IC neuron (20108; CF = 3.4 kHz; threshold = -65 dB re maximum system output) that was not classified as ITD-sensitive to either SAM or transposed tones. The Figure follows the same format as Figure 4. (B) PSTHs to 200 ms stimulation (C) – (H) Raster plots for recordings at 10, 20, 40, 80, 160 and 320 Hz f_m respectively. Raster plots are not displayed for the responses at 640 Hz f_m which are included in (I). Black bars indicating the middle 350 ms of the stimulus are not shown.

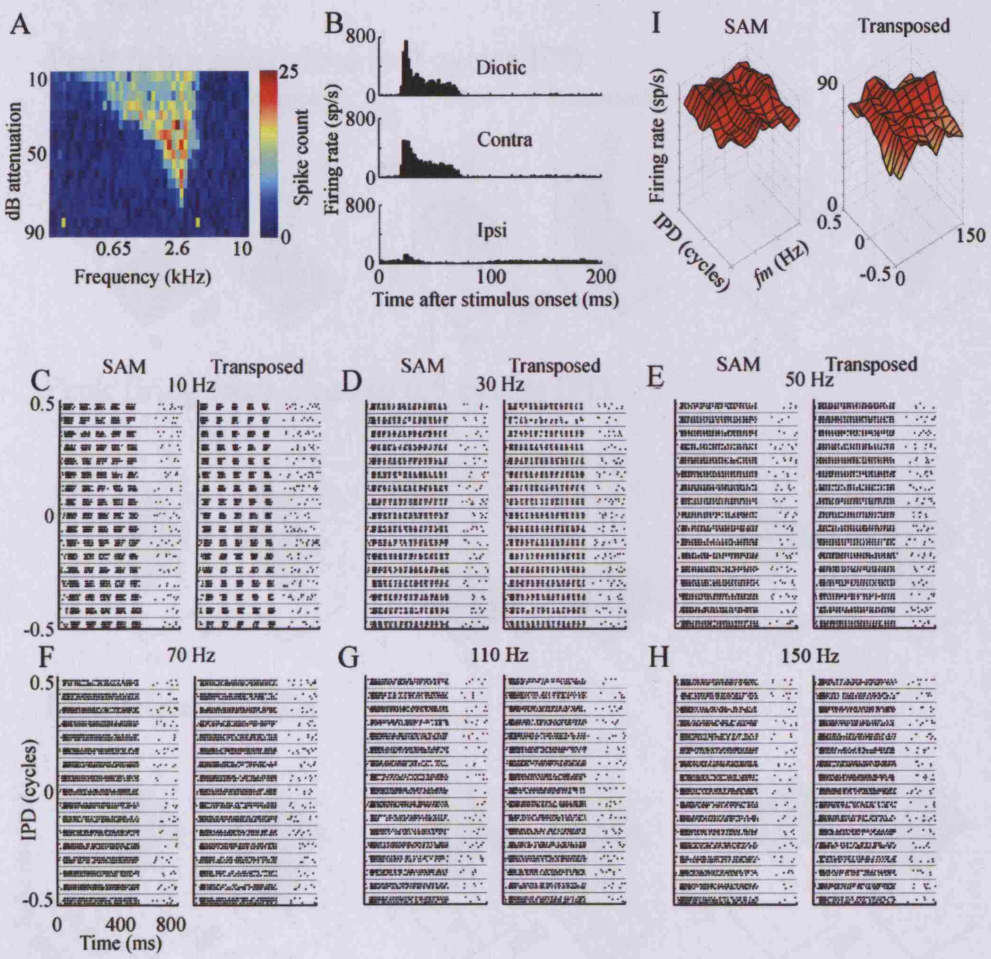

Figure 3.4. The responses of an IC neuron (25805; CF = 2.8 kHz; threshold = -79 dB re maximum system output) that was not classified as ITD-sensitive to either SAM or transposed tones. The figure follows the same format as Figure 4. (C) – (H) Raster plots for recordings at 10, 30, 50, 70, 110 and 150 Hz f_m respectively. Raster plots are not displayed for the responses at 90 or 130 Hz f_m which are included in (I). Black bars indicating the middle 350 ms of the stimulus are not shown.

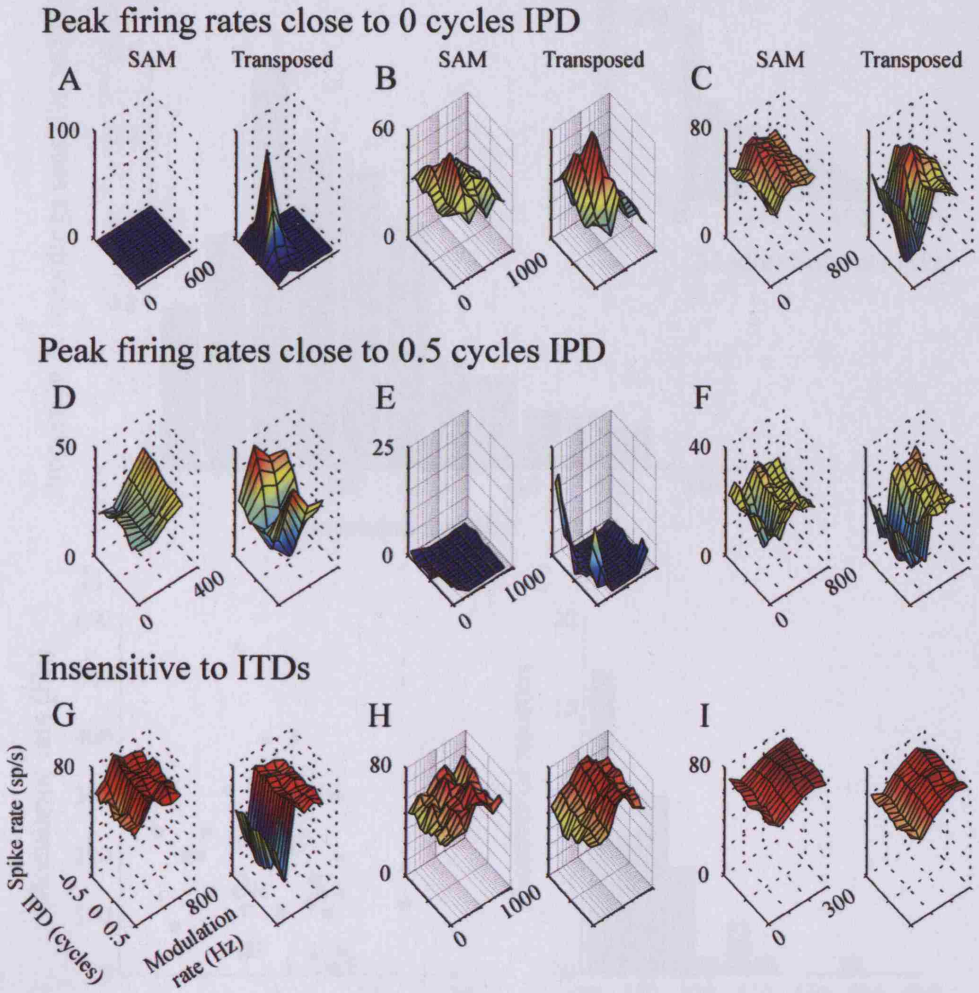

Figure 3.5. Firing rates in response to ITDs conveyed by SAM (left) and transposed (right) tones, at different f_m are shown. 3D mesh plots are presented in the same format as in Figure 4I, for 9 IC neurons. Top row: ITD-sensitive neurons with peak responses close to 0-μs ITD. (A) Neuron 25806, CF = 7.8 kHz, threshold = -44 dB re maximum system output. (B) Neuron 12505, CF = 3.6 kHz, threshold = -81 dB re maximum system output. (C) Neuron 19202, CF = 2 kHz, threshold = -52 dB re maximum system output. Middle row: ITD-sensitive neurons with a trough in response close to 0-μs ITD. (D) Neuron 27204, CF = 2.2 kHz, threshold = -55 dB re maximum system output. (E) Neuron 12821, CF = 8.8 kHz, threshold = -62 dB re maximum system output. (F) Neuron 20818, CF = 4.8 kHz, threshold = -70 dB re maximum system output. Bottom row: Neurons classified as insensitive to ITDs. (G) Neuron 22507, CF = 5.1 kHz, -85 dB re maximum system output. (H) Neuron 20804, 5.4 kHz, -81 dB re maximum system output. (I) Neuron 25605, CF = 3.8, threshold = -73 dB re maximum system output.

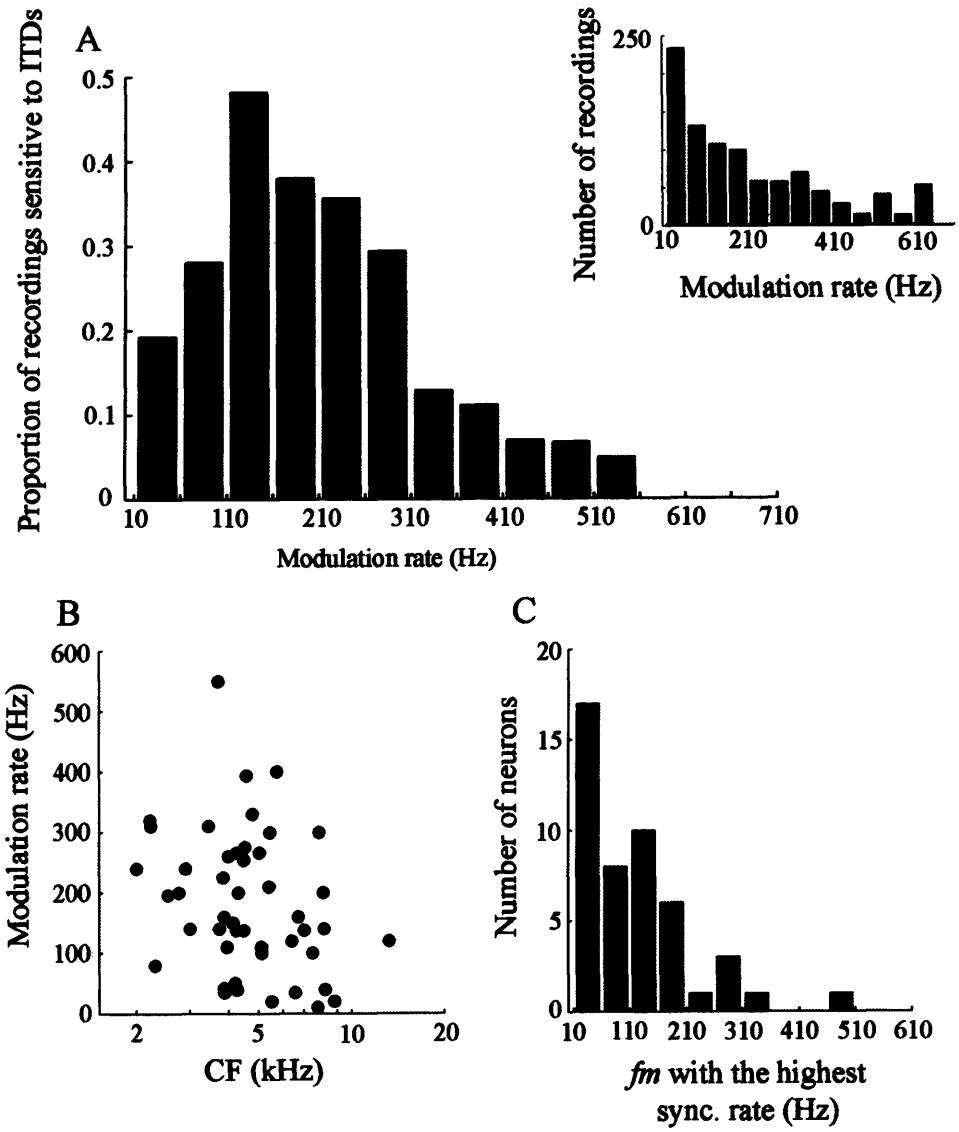

Figure 3.6. Sensitivity to ITDs as a function of frequency. A) The proportion of recordings that were considered sensitive to ITDs plotted as a function of modulation rate. Inset: The total number of recordings at each modulation rate. B) The maximum f_m at which sensitivity to ITDs occurred as a function of the CF of each neuron. C) A histogram showing the distribution of the f_m at which recordings had the highest synchronization.

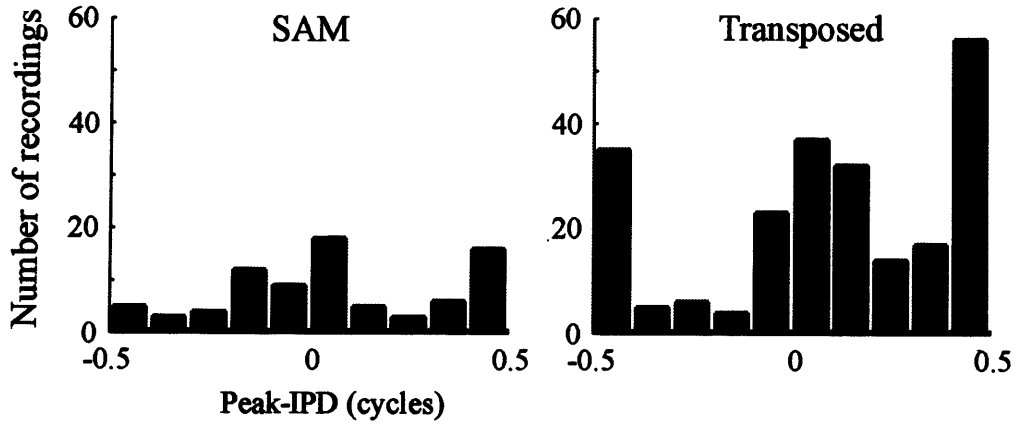

Figure 3.7. Peak-IPDs of all ITD-sensitive recordings in response to SAM (left) and transposed (right) tones.

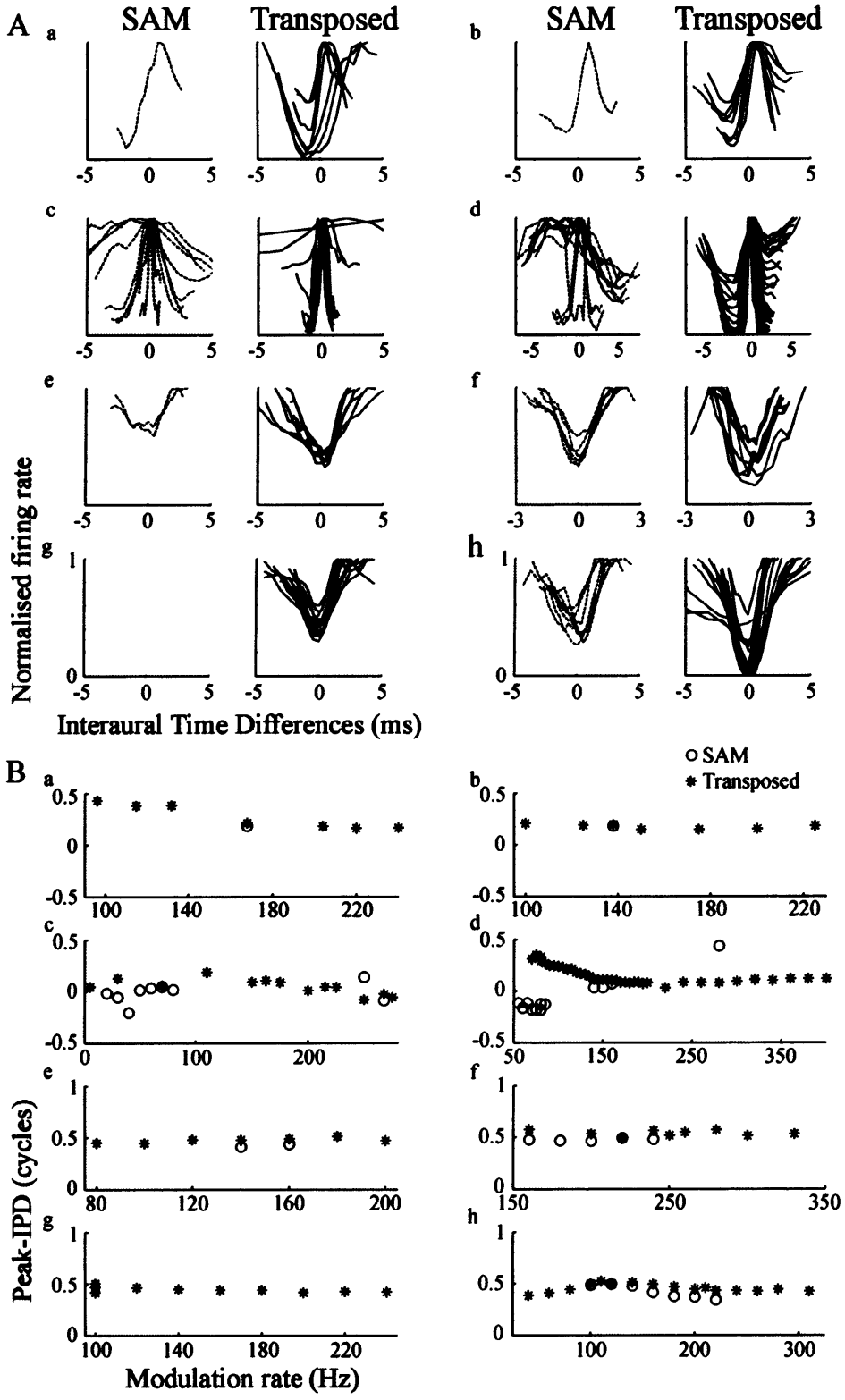

Figure 3.8. See next page for legend.

Figure 3.8. ITD-tuning as a function of modulation rate. (A) Normalized firing rates, at different f_m, plotted as a function of ITD for 8 IC neurons (a)-(h) in response to transposed (left of each pair) and SAM (right of each pair) tones. Neurons: (a) 19202, (b) 22207, (c) 18808, (d) 20709, (e) 22205, (f) 20818, (g) 20810 & (h) 27204. (B) Peak-IPDs plotted as a function of modulation rate for the same 8 neurons in (A) in response to transposed (closed circles) and SAM (open circles) tones. Note that the y-axes in a-d and in e-h are different.

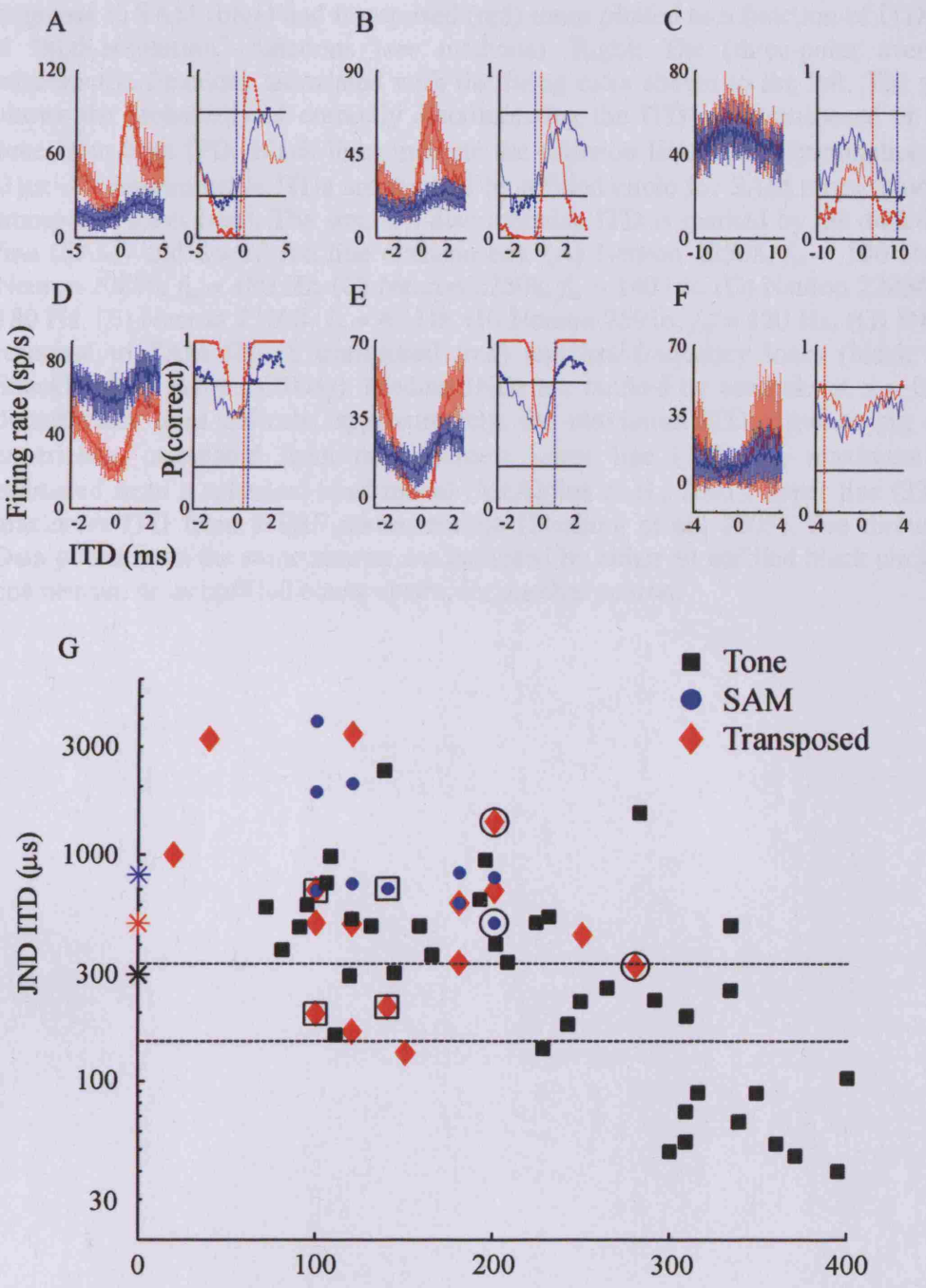

Figure 3.9. See next page for legend.

Figure 3.9.The derivation of neural ITD discrimination thresholds (just noticeable differences, JNDs). (A)-(F) Left: The mean firing rates and standard deviations in

response to SAM (blue) and transposed (red) tones plotted as a function of ITD from 6 "high-resolution" functions (see methods). Right: The (three-point averaged) neurometric functions associated with the firing rates shown to the left. The y-axis shows the probability of correctly discriminating the ITD of a transposed or SAM tone from 0-μs ITD. Black lines indicate the criterion level for discrimination from 0 μs. All discriminable ITDs are marked by a filled circle for SAM tones (blue) and transposed tones (red). The smallest discriminable ITD is marked by the dotted blue line (SAM) and dotted red line (transposed). (A) Neuron 22508, f_m = 100 Hz. (B) Neuron 20810, f_m = 180 Hz. (C) Neuron 22508, f_m = 140 Hz. (D) Neuron 27204, f_m = 180 Hz. (E) Neuron 23008, f_m = 40 Hz. (F) Neuron 25916, f_m = 120 Hz. (G) JNDs in response to SAM (blue), transposed (red) and low-frequency tones (black; from Shackleton & Palmer (2004)). Median JNDs are marked by asterisks at x = 0. The dotted black lines indicate, approximately, the maximum ITD a guinea pig could experience, calculated from two sources; lower line (150 μs), maximum ITD estimated from a spherical head model (McAlpine et al., 2001), upper line (330μs), maximum ITD from HTRF measurements (Sterbing et al., 2003). See discussion. Data points from the same neuron are indicated by either an unfilled black circle, for one neuron, or an unfilled black square, for another neuron.

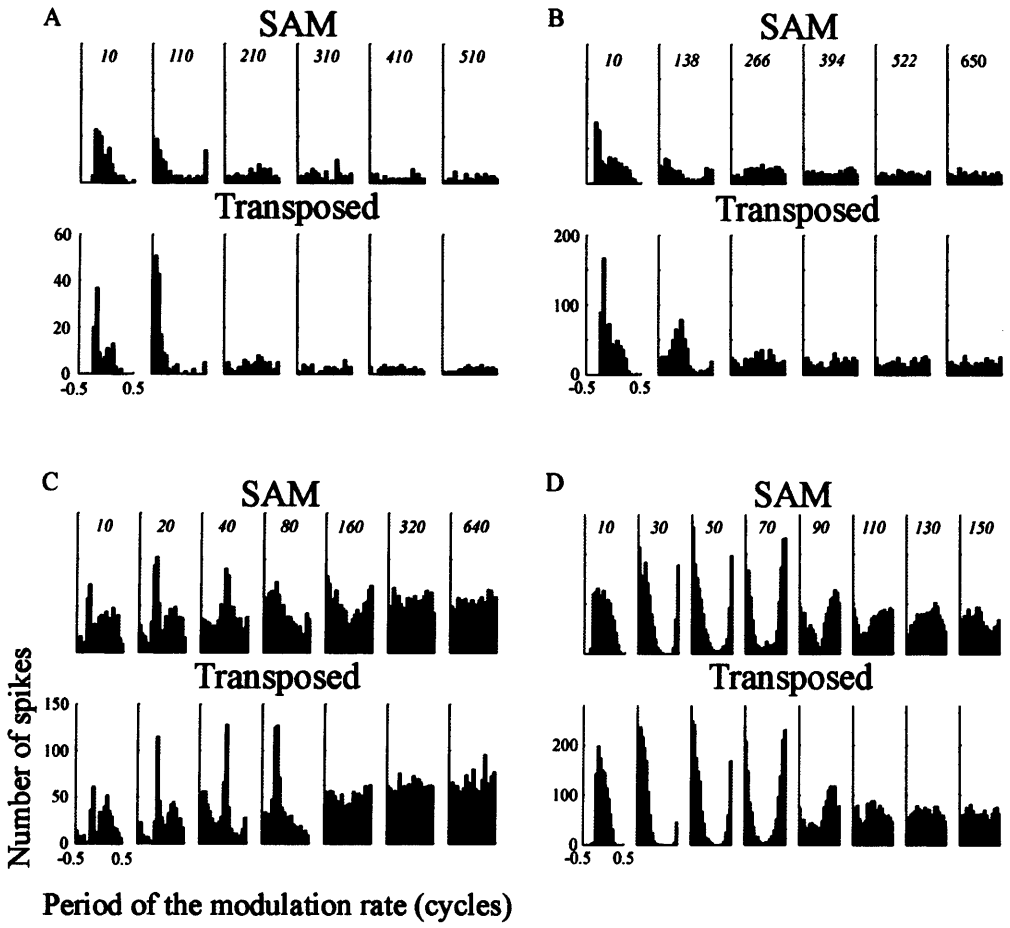

Figure 3.10. Phase locking as a function of modulation rate. (A) Period histograms of spikes grouped over one cycle of modulation rate in response to SAM (top) and transposed (bottom) tones for the neuron in Figure 3.1.2.1. The corresponding modulation rates (f_m), in Hz, are marked in italics atop the histograms of responses to SAM tones. (B) Period histograms for the neuron in Figure 3.1.2.2, (C) Period histograms for the neuron in 3.1.2.3 and (D) Period histograms for the neuron in Figure 3.1.2.4.

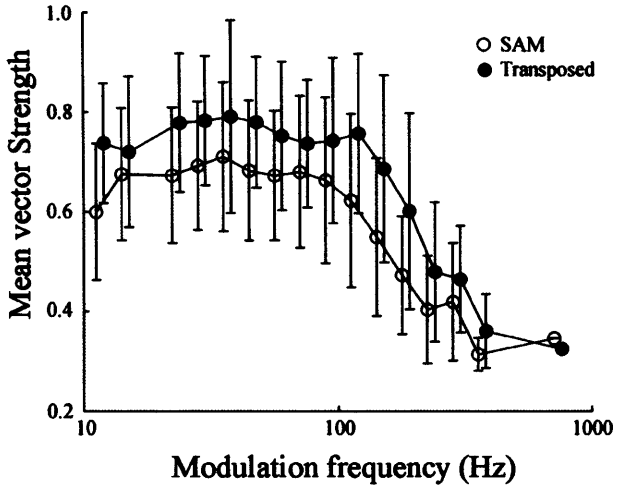

Figure 3.11. Phase locking as a function of modulation rate. The mean and standard deviation vector strengths of phase locking, from all neurons, to SAM (open circles) and transposed (closed circles) tones are plotted with error bars showing ± 1 standard deviation of the mean. The vector strengths obtained with the transposed tones have been slightly offset horizontally.

4. Characterising ITD-sensitivity in response to SAM and transposed tones

4.1. Introduction

In Chapter 3, neural firing rates were described in response to SAM and transposed tones with different f_m and ITDs. More neurons were sensitive to ITDs in response to transposed as compared to SAM tones. The degree of sensitivity to ITDs was also assessed by a discrimination analysis (see Section 3.3.3) and was found to be greater in response to transposed tones as compared to SAM tones. In this Chapter, the firing rates of ITD-sensitive recordings, as a function of ITD, will also be referred to as ITD-tuning functions. Lower JNDs were obtained in response to transposed tones because firing rates changed more rapidly as the ITD changed from 0 μs than in response to SAM tones i.e. the gradients of the ITD-tuning functions were steeper around 0-μs ITD in response to transposed as compared to SAM tones. This could occur because ITD-tuning functions had higher modulation depths in response to transposed as compared to SAM tones or because ITD-tuning functions were narrower around their peak or trough in response to transposed tones. In many of the examples in Chapter 3, ITD-sensitive recordings were described as more modulated in response to transposed tones as compared to SAM tones. However, the modulation depths were not quantified. In the first results Section in this Chapter (Section 4.3.1), ITD-sensitive recordings are further characterised to investigate the differences which underlie greater sensitivity to ITDs in transposed as compared to SAM tones. The steepest slopes (maximum change in firing rate as a function of ITD), depths of modulation, mean firing rates and half-widths (widths at half-way between their maximum and minimum firing rates) of ITD-sensitive recordings in response to SAM, transposed and pure tones are compared.

In the second results Section of this Chapter (Section 4.3.2), model ITD-tuning functions in response to SAM and transposed tones are presented. Monaural

responses were constructed by employing the same model of peripheral processing used to account for human ITD-discrimination thresholds (Bernstein & Trahiotis, 2002). This was followed by calculating cross-products of the modelled monaural responses to simulate coincidence detection. The modelled ITD-tuning functions are contrasted with the characteristics of neural ITD-tuning functions described in Section 4.3.1. Components of the model of peripheral processing were altered to investigate the how peripheral processing might influence the shape of neural ITD-tuning functions.

4.2. Data Analysis

Data from low-CF neurons

Firing rates obtained from low-CF neurons in response to tones with IPDs between ± 0.5 cycles were examined from the recent data archive of the laboratory. Responses were analysed from 61 low-CF neurons from a total of 37 animals. They included response to presentations of either 50 or 100 ms tones, ramped on for 2 ms or 10 ms with a $\cos^2$ gate. There was no onset-ITD. The tone frequencies were equal to the CF of the neuron and had intensities between 20 and 30 dB above threshold. Only responses with Rayleigh coefficient > 13.815 ($p < 0.001$; Mardia & Jupp, 1999) were considered ITD-sensitive.

Steepest slopes of ITD-tuning functions

The gradients of ITD-tuning functions were estimated by first fitting a linear regression of every three consecutive points on the 3-point smoothed ITD-tuning functions. The maximum and minimum gradients were identified, representing the steepest positive and steepest negative slopes of the ITD-tuning function. The *steepest slope* was defined as the maximum of the steepest positive and negative slopes.

Half-widths of ITD-tuning functions

The half-width was defined as the width of the peak of ITD-tuning functions at half-way between the maximum and minimum firing rates. Half-widths were estimated from ITD-functions smoothed by 3-point averaging. Linear interpolation was used to determine the two IPDs half-way between the maximum and minimum firing rate. The half-width, in cycles, was the difference between these two values.

Box and whisker plots

Box and whisker plots were constructed as follows. The median was marked by a red circle, the inter-quartile range was marked by a thick black bar or box and the range was marked by a thin black line. Outliers were marked with red crosses. Outliers were classified as such if they were outside the inter-quartile range by more than 1.5 times the inter-quartile range. If there were fewer than 5 data points in the frequency bin, only the median and range of the data are shown.

4.3. Results

4.3.1. Characterisation of ITD-sensitive recordings in response to SAM, transposed and low-frequency pure tones.

In this Section, ITD-tuning functions in response to SAM and transposed tones are characterised as described in section 4.1. In addition, ITD-tuning functions in response to pure tones, previously obtained in the laboratory, are characterised and compared to ITD-tuning functions in response to SAM and transposed tones.

The distribution of the steepest slopes of ITD-sensitive recordings is shown in box and whisker plots in Figure 4.1. Data were divided into six frequency bins according

to the f_m of the high-frequency stimulus or the frequency of the pure tone stimulus. Table 2 shows the number of recordings in each frequency bin. Consistent with the lowest neural JNDs for ITD-discrimination being obtained from responses to transposed tones as compared to SAM tones, the steepest slopes of the ITD-tuning functions in response to transposed tones could be higher than in response to SAM tones, particularly for f_m of 40 - 320 Hz.

The magnitude of the slope of an ITD-tuning function is partly dependent on its modulation depth (the difference between the maximum and minimum firing rate). The modulation depths of all ITD-tuning functions in response to SAM and transposed tones (as well as pure tones) are displayed in Figure 4.2, in the same format as in Figure 4.1. The distribution of modulation depths was similar to that of the steepest slopes (compare Figures 4.1 & 4.2). A 3-factor ANOVA incorporating the main effects of the neuron, f_m and stimulus type provided statistical evidence that all three factors influenced the modulation depths (all $p < 0.01$). The proportion of ITD-tuning functions with modulation depths above 30 spikes per second (sp/s) was small in response to SAM as compared to transposed tones [7.4 % (6/81) in response to SAM tones and 42 % (96/229) in response to transposed tones]. In response to transposed tones with f_m between 80 Hz and 160 Hz, 49/84 (58%) of ITD-tuning functions had modulation depths > 30 sp/s. This f_m range had the greatest proportion of ITD-tuning functions with modulation depths > 30 sp/s and was also the f_m range where most recordings were sensitive to ITDs in transposed tones (see also Figure 3.6A and Table 2).

The sample of responses to low-frequency tones was small and biased towards frequencies above 160 Hz (Table 2). At f_m above 160 Hz human sensitivity to ITDs in the envelope of high-frequency stimuli decreases, whereas sensitivity to ITDs in pure tones increases (see Figure 3.6; Bernstein & Trahiotis, 2002). Consistent with this, greater modulation depths and greater steepest slopes were observed in response to ITDs in pure tones above 160 Hz than in response to SAM and transposed tones above 160 Hz f_m (Figures 4.1 & 4.2; note the different scale of the ordinates in C

compared to A & B). 96 % (25/26) of ITD-tuning functions in response to pure tones had modulation depths > 30 sp/s. It is not clear whether ITD-tuning functions in response to pure tones below 160 Hz would also have greater modulation depths than ITD-tuning functions in response to SAM and transposed tones with f_m below 160 Hz. The discrimination analysis presented in Section 3.3.3 provides a fairer comparison between the sensitivity to ITDs in response to low-frequency tones and high-frequency stimuli.

It might be predicted that the larger modulation depth of ITD-tuning functions in response to transposed tones is accompanied by increased mean firing rates, as compared to responses to SAM tones. In some of the examples given in Chapter 3 (see Figures 3.1, 3.2, 3.5 & 3.9), the higher modulation depth in response to transposed tones occurred by means of both an increase in maximum firing rate and a decrease in minimum firing rate, compared to the response to SAM tones. In these cases, the mean firing rates of ITD-tuning functions in response to transposed and SAM tones were similar. The mean firing rates of all recordings were calculated and compared in response to SAM and transposed tones.

Recordings were always obtained in response to transposed and SAM tones at the same f_m. For each of these paired recordings, the difference in mean firing rate in response to transposed tones and SAM tones is displayed in the histogram in Figure 4.3A. The differences in mean firing rate were distributed around 0 sp/s. There was no neural correlate of the increased sensitivity to ITDs in transposed tones in the mean firing rates of this sample of responses from the IC [a multi-factor ANOVA for the main effects of the neuron, the f_m and the stimulus (Transposed or SAM) on the mean firing rates returned $p = 0.27$ for the stimulus].

Although the differences in mean firing rates were distributed around 0 sp/s, some paired recordings had large (> ±20 sp/s) differences in the firing rates in response to SAM and transposed tones. Figure 4.3B shows a histogram of differences between mean firing rates in response to transposed and SAM tones, from paired recordings

that were sensitive to ITDs in one or both of the stimuli. In this sub-population of the data shown in Figure 4.3A, a greater proportion of recordings had large (> ±20 sp/s) differences in mean firing between SAM and transposed tones. However, it was still not the case that firing rates were always higher (or lower) in response to transposed tones. These results are consistent with the finding that increased modulation depths of ITD-tuning functions in response to transposed tones are not solely caused by increases in maximum firing rates compared to the responses to SAM tones.

The histogram in Figures 4.3C shows the differences in mean firing rates in response to transposed and SAM tones for a subset of the data in Figure 4.3B that were ITD-sensitive and had modulation depths > 30 sp/s in response to either stimulus. In these cases, there were more recordings with higher mean firing rates in response to transposed tones than SAM tones. This suggests that for neurons and f_m at which firing rates are most strongly modulated as a function of ITD, firing rates tend to be higher in response to transposed tones than SAM tones. However, this was not exclusively the case.

The final measurement made to characterise ITD-tuning functions was the half-width. It has been suggested that half-widths become more narrow in the ascending auditory pathway and that sharpening of tuning to ITDs represents greater sensitivity to ITDs in single neurons (Fitzpatrick *et al.*, 2000). It might be predicted that half-widths of ITD-tuning functions were sharper in response to transposed as compared to SAM tones. However, the half-widths of ITD-tuning functions were not significantly influenced by the stimulus type [multi-factor ANOVA examining the main effects of the neuron, f_m and stimulus type (Transposed or SAM): $p = 0.27$, $p = 0.99$ and 0.12 respectively]. Figure 4.4 displays the half-widths of paired recordings that were ITD-sensitive in response to both SAM and transposed tones. There was a trend (which was not significant) for ITD-tuning functions in response to transposed tones to have smaller half-widths than in response to SAM tones (Median halfwidths = 0.45 vs. 0.48, Wilcoxon paired signed-rank test; $z = 1.87$, $p = 0.06$). These data suggest that a narrower half-width is not necessarily required to underlie greater sensitivity to ITDs.

Histograms of the half-widths of all ITD-tuning functions are shown in Figure 4.5 (clear bars). Particularly in response to transposed tones, the half-widths fell into two clusters, centred below and above 0.5 cycles. Batra *et al.* (1993) found half-widths were dependent on whether ITD-tuning functions in response to SAM tones were dominated by a peak or trough in firing rate (i.e. "peak-type" or "trough-type", see Section 3.3.2). ITD-tuning functions that had peak-IPDs close to 0.0 (-0.1 to 0.3 cycles) will be referred to as "peak-type" and ITD-tuning functions that had close to 0.5 cycles (<0.4 cycles or >0.3 cycles) will be referred to as "trough-type". The half-widths of "peak-type" and "trough-type" ITD-tuning functions are shown in Figure 4.5. In response to transposed tones, the median half-with of "peak-type" ITD-tuning functions was 0.25 cycles whereas the median half-width of "trough-type" ITD-tuning functions was 0.56 cycles. The median half-width around the trough of "trough-type" ITD-tuning functions was 0.44 cycles, still wider than the half-widths of "peak-type" ITD-tuning functions and consistent with Batra *et al.* (1993). The bimodal distribution of half-widths is correlated with the bimodal distribution of peak-IPDs (see Figure 3.7).

4.3.2. Modelled responses to SAM and transposed tones

The most consistent differences found between ITD-tuning functions in response to SAM and transposed tones were the steeper slopes and greater modulation depths in response to transposed tones. Bernstein & Trahiotis (2002) successfully account for human sensitivity to ITDs in SAM and transposed tones using a model of peripheral processing followed by calculation of the normalised interaural correlation. The same peripheral model [incorporating gamma-tone filtering, half-wave square law rectification, compression and low-pass filtering (see Section 2.2.2)] was used to simulate monaural responses to SAM and transposed tones. Cross-products were then calculated to create modelled ITD-tuning functions and these were compared to the neural recordings obtained in response to SAM and transposed tones. The parameters

of the model of peripheral processing were varied to understand how peripheral processing could affect the shape of model ITD-tuning functions.

Modelled monaural responses to SAM and transposed tones with f_m of 80 Hz and 160 Hz and with f_c of 4 kHz are shown in Figures 4.6A & B. The modelled monaural responses maintained periodicity at the f_m but were smoothed and attenuated by the filtering steps. In response to transposed tones (red), the modelled monaural responses no longer showed the "off-periods" present in transposed tones (see Figure 1.7.1). However, they did show greater depths of modulation and were lower at their minimum compared to model responses derived from SAM tones (black). The reduction in modulation depth at f_m = 160 Hz as compared to 80 Hz is caused by the low-pass filter, which provides an upper limit on the f_m at which modelled sensitivity to ITDs occurs.

Phase-locking to the f_m in the modelled monaural responses was compared to phase-locking to the f_m in the neural responses to binaural stimuli shown in Figure 3.11. Vector strengths of model monaural responses to SAM and transposed tones with f_m = 80 Hz were 0.26 and 0.55, respectively. This is lower than the mean vector strengths of phase-locking to the f_m observed in neural recordings from the IC (Figure 3.11), and might partly be explained by an improvement in phase-locking between ANFs and coincidence detector neurons not incorporated in the model (Joris *et al.*, 1994b; Louage *et al.*, 2005). This difference might also be explained by a higher "dc component" in the modelled monaural responses. Half-widths of period histograms created from the modelled monaural firing rates (not shown) were similar (around 0.51 cycles) for waveforms derived from both SAM and transposed tones with f_m = 160 Hz but different when the f_m = 80 Hz (0.58 cycles for SAM tones and 0.46 cycles for transposed tones). The period histograms constructed from neural spike times, and shown in Figure 3.10, were often more narrow in response to transposed compared to SAM tones, particularly at the lower f_m.

The cross-products of the waveforms were calculated following peripheral processing to simulate the ITD-tuning function of a single coincidence detector neuron (Figures 4.6C & D). The abscissa in Figures 4.6C & D represents the IPDs (in cycles of the f_m) in the signal and the cross-product represents the firing rate of the single neuron at each IPD in the stimulus. The modulation depths of the modelled ITD-tuning functions were lower when derived from SAM as compared to transposed tones, consistent with the neural ITD-tuning functions. The modulation depths were also lower for f_m =160 Hz compared to f_m = 80 Hz due to the low-pass filter in the model of peripheral processing. This trend was present in the maximum modulation depths obtained from neural ITD-tuning functions (Figure 4.2) but was not necessarily present within the responses of single neurons (Chapter 3).

The differences in mean firing rates of neural ITD-tuning functions in response to the SAM and transposed tones were distributed around 0 sp/s (see Figure 4.3). In contrast to this, the mean cross-product was higher in response to SAM as compared to transposed tones. This can partly be explained due to the higher *rms* intensity of the SAM tones; stimuli with equal maximum intensity were modelled because neural firing rates were also obtained in response to stimuli with equal maximum intensity. The mean cross-product was also higher in response to SAM tones because the filtering in the peripheral model attenuates more of the spectral components of a transposed tone as compared to an SAM tone. This is due to the wider spread of the spectral components in a transposed as compared to an SAM tone (Figure 1.1).

Bernstein & Trahiotis (2002) accounted for human ITD-discrimination thresholds by calculating the normalised correlation (shown in Figures 4.6E&F). This can be considered as the cross-product normalised to a maximum of 1. Human ITD-discrimination thresholds were estimated by assuming that listeners were sensitive to a constant level of decorrelation of the stimulus. The normalised correlation fell more quickly from its maximum at 0-cycles IPD, as IPDs change from 0-cycles IPD towards positive (or negative) IPDs, for the transposed as compared to the SAM stimulus. This correctly predicted that thresholds for detection of ITDs are smaller in

response to transposed tones as compared to SAM tones because a constant level of the normalised correlation (any value < 1) is achieved with a smaller change in IPD from cycles.

The half-widths of the modelled ITD-tuning functions were almost identical in response to SAM and transposed tones. This is easily observed in Figures 4.6G & H in which the model ITD-tuning functions are normalised between 0 and 1. The half-widths are indistinguishable in response to SAM and transposed tones on the scale used. This seemed counter-intuitive given the steeper rise time and the shorter "on-period" of transposed as compared to SAM tones. However, this was consistent with the half-widths of neural ITD-tuning functions which were also similar in response to SAM and transposed tones. The model of peripheral processing was altered to explore the origin of the striking similarity in the shape of modelled ITD-tuning functions in response to SAM and transposed tones.

Figure 4.7 shows modelled responses to SAM and transposed tones in the same format as Figure 4.6 but the data were obtaining using a "gamma-tone only" model of peripheral processing. This model generated monaural responses simply from the Hilbert envelope of the stimulus, after filtering with a gamma-tone filter (4.7A & B). The gamma-tone filter smoothed the waveforms and reduced the "off-period" of the transposed waveform. This was particularly noticeable at f_m = 160 Hz as compared to 80 Hz and occurred because the spectral components of a 160 Hz transposed tone are more widely spaced and therefore more heavily attenuated by the gamma-tone filter.

The modulation depths of the modelled ITD-tuning functions, using the "gamma-tone only" were almost identical in response to SAM and transposed tones (Figure 4.7C & D). However, the normalised correlation derived from transposed tones fell more rapidly from 1, as the IPD increased (or decreased) from 0-cycles IPD, than the normalised correlation derived from SAM tones (Figure 4.7E & F). This occurred because the cross-products, despite having similar modulation depths, fell closer to zero when derived from the transposed as compared to the SAM waveform. This, in

turn, occurs because the modelled monaural responses derived from transposed tones fell closer to zero and remained low for a longer duration than the responses derived from SAM tones. The differences in the normalised correlation derived from SAM and transposed waveforms were small compared to that achieved by the model of peripheral processing used by Bernstein & Trahiotis (compare Figures 4.6E & F and 4.7E & F) and are not sufficient to account for human psychophysical performance. The half-widths of the modelled ITD-tuning function in response to transposed tones were narrower than the half-width of the modelled ITD-tuning function in response to SAM tones, particularly at f_m = 80 Hz (Figure 4.7G & H).

The results in Figure 4.7 show that a "gamma-tone only" model of peripheral processing that simply derives the envelope of a SAM and transposed waveform is insufficient to account for the similar half-widths but different modulation depths of neural ITD-tuning functions. The compression and low-pass filtering were systematically altered in the peripheral model used by Bernstein & Trahiotis (2002). It was found that both the compression and low-pass filtering affected the half-widths and modulation depths of the modelled ITD-tuning functions.

The result of omitting the compression from the peripheral model used by Bernstein & Trahiotis is shown in Figure 4.8. The modelled monaural responses derived from stimuli with f_m = 160 Hz were smoothed and had similar shapes (Figure 4.8B). However, the waveforms with 80 Hz f_m differed in shape in response to SAM and transposed tones when compression was omitted, particularly at their lowest values (Figure 4.8A). Both the gamma-tone filtering and the low-pass filtering affected the stimuli with 160 Hz f_m greater than the stimuli with 80 Hz f_m. The modulation depth of the modelled ITD-tuning function was actually higher when derived from *SAM* as compared to transposed waveforms (Figures 4.8C & D). In the absence of compression, low-pass filtering reduced the modulation depth more in response to transposed tones as compared to the SAM tones because transposed tones have more spectral components that are heavily attenuated by the filtering in the model. Compression effectively compensates for the greater reduction in the modulation

depth caused by low-pass filtering of the transposed tones by reducing the modulation depth more in response to SAM as compared to transposed tones.

The half-widths of the model ITD-tuning functions, in the absence of compression, were almost identical when the f_m was 160 Hz but were narrower for transposed tones when the f_m was 80 Hz (Figure 4.8G & H). The gamma-tone filter, half-wave rectification and low-pass filter were sufficient to result in equal half-widths in response to SAM and transposed tones at f_m = 160 Hz. However, compression was required to obtain equal half-widths in response to SAM and transposed tones with f_m = 80 Hz. It was found that the half-width of model ITD-tuning functions could be controlled by altering the magnitude of the compression. The half-widths of model ITD-tuning functions derived from transposed tones increased in relation to those derived from SAM tones as the compression increased (i.e. as the exponent to which the waveform was raised was increased). Both low-pass filtering and compression, in the peripheral model used by Bernstein & Trahiotis (2002), are required to account for the almost identical half-widths but higher modulation depths of model ITD-tuning functions in response to transposed as compared to SAM tones. The critical role of compression in their peripheral model, for accounting for human psychophysical thresholds, was noted by Bernstein & Trahiotis (2002). The difference in the normalised correlation, between the waveforms derived from SAM and transposed tones, was least in the model in which compression was omitted (Figure 4.8E & F).

4.4. Discussion

The modulation depth of ITD-tuning functions was found to be greater in response to transposed as compared to SAM tones. The half-widths and mean firing rates could differ in response to SAM and transposed tones but were not consistently higher or lower in response to either stimulus. The higher modulation depths, but similar half-widths of ITD-tuning functions in response to transposed as compared to SAM tones,

is consistent with the higher maximum slopes observed in response to transposed tones and with the greater sensitivity to ITDs in transposed tones reported in Chapter 3.

The model of peripheral processing used by Bernstein & Trahiotis (2002), followed by cross-correlation of the resulting waveforms, predicts that ITD-tuning functions in responses to SAM and transposed tones differ primarily in their modulation depths. This was consistent with the physiological results. The modulation depths of modelled ITD-tuning functions were dependent on the nature of the model of peripheral processing. The modelling suggests that compression was required to account for the higher modulation depths in response to transposed as compared to SAM tones. Both the filtering and compression in the peripheral model were responsible for the almost identical half-widths of modelled ITD-tuning functions in response to SAM and transposed tones.

The half-widths of neural ITD-tuning functions were similar in response to SAM and transposed tones but, in contrast to the modelled ITD-tuning functions, half-widths could vary considerably between neurons. This variation was partly due to the mixed sample of neurons that might be described as "peak-type" and had narrower half-widths than those that might be described as "trough-type" (see Figure 4.5; Batra *et al.* 1993). In addition, it seems likely that neural mechanisms not included in model influence the half-width of ITD-tuning functions e.g. the convergence of inputs with different binaural response properties in the IC (McAlpine *et al.*, 1998; Fitzpatrick *et al.*, 2000; Shackleton *et al.*, 2000).

Cross-products of the model ITD-tuning functions to SAM and transposed tones were higher in response to SAM tones as compared to transposed tones. There was no evidence in the neural data to suggest that mean firing rates in response to SAM tones were always higher than in response to transposed tones. Neural mechanisms not included in the model, such as the intrinsic response properties of different neurons (Frisina, 2001) or the timing of excitatory and inhibitory inputs to coincidence

detector neurons (Grothe & Park, 2000), could cause differential firing rates to SAM and transposed tones. These processes can also cause different neurons to exhibit different rMTFs (see Section I.1.4) and different ITD-sensitivity to the same f_m (see Chapter 3), features which were not captured by the models used in this Chapter.

The model of peripheral processing used in this Chapter was developed to account for ITD-discrimination thresholds obtained from human listeners (Bernstein & Trahiotis; 2002). Therefore, the parameters of the filtering, half-wave rectification and compression were not optimised for processing that occurs in the auditory periphery of the guinea pig. In addition, known physiological processes were not included in the model e.g. the peripheral processing did not include adaptation that occurs in ANFs and the role of inhibitory inputs onto coincidence detector neurons was not modelled. Modelling the output of coincidence detector neurons by calculating cross-products or cross-correlations is an over-simplification of the process that occurs in these neurons (Batra & Yin, 2004). Despite these caveats, the modelling does provide insight into how compression and filtering might affect the shape of ITD-tuning functions. It provides an example of how neural ITD-tuning functions in response to SAM and transposed tones might result in similar half-widths but different modulation depths.

In summary, ITD-tuning functions in response to SAM and transposed tones primarily differ in their modulation depths, which are greater in response to transposed tones. The half-widths of ITD-tuning functions in response to SAM and transposed tones are similar and this can be predicted by employing the same model of peripheral processing used to account for human psychophysical data. Alterations of the peripheral model demonstrate the importance of compression and filtering both in predicting the similar half-widths and different modulation depths of ITD-tuning functions in response to SAM and transposed tones, and in accounting for psychophysical thresholds.

4.5. Figures and Tables for Chapter 4

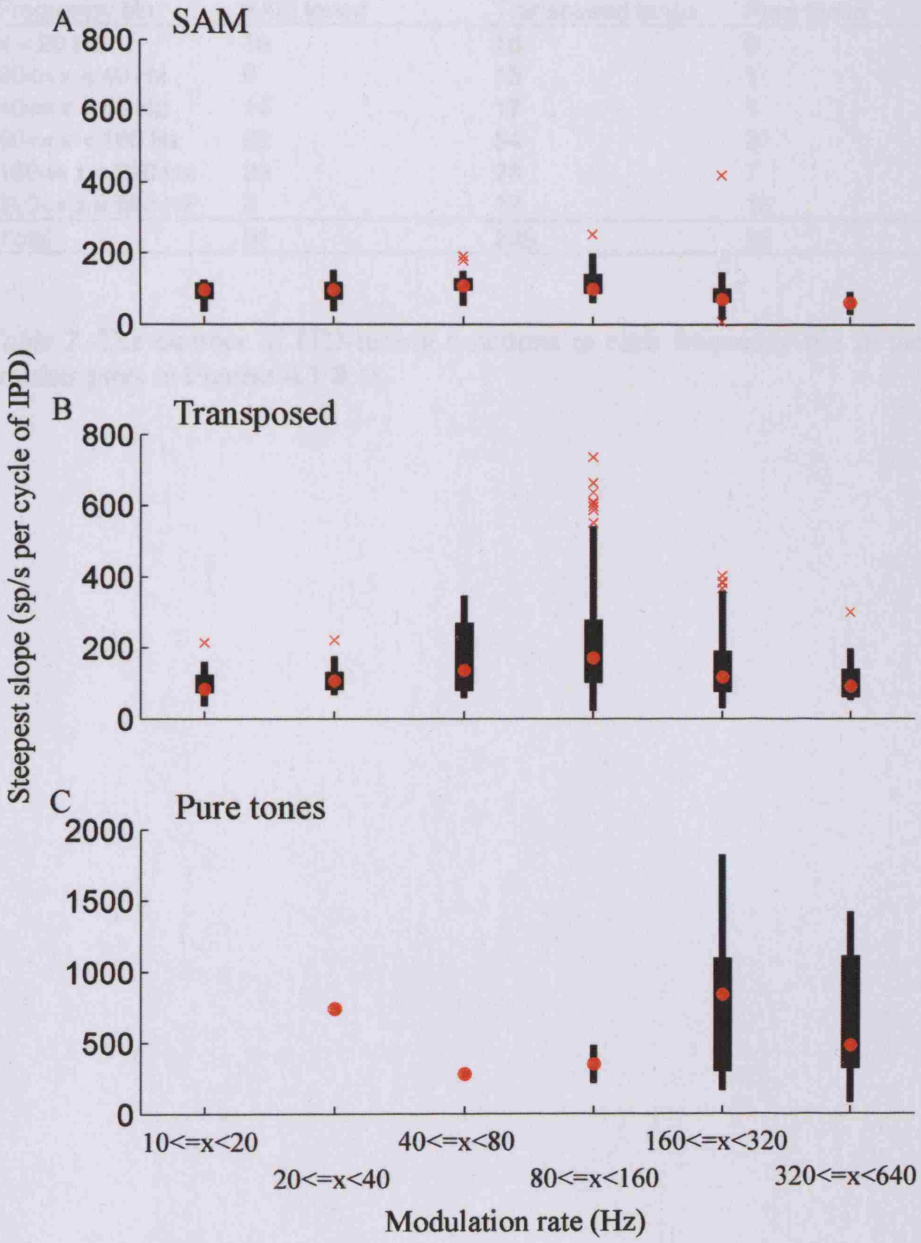

Figure 4.1. Box and whisker plots showing the distribution of the steepest slopes of all ITD-sensitive recordings, grouped according to the frequency of the stimulus. The median (red circle), inter-quartile range (thick black bar or box) and range (thin black line) are shown along with any outliers (red crosses). See Section 4.2 for the definition of outliers.

Frequency bin	SAM tones	Transposed tones	Pure tones
x < 20 Hz	10	18	0
20<= x < 40 Hz	9	15	1
40<= x < 80 Hz	14	17	1
80<= x < 160 Hz	22	84	2
160<= x < 320 Hz	23	78	7
320<= x < 640 Hz	3	17	15
Total	81	229	26

Table 2. The number of ITD-tuning functions in each frequency bin in the box and whisker plots in Figures 4.1 & 2.

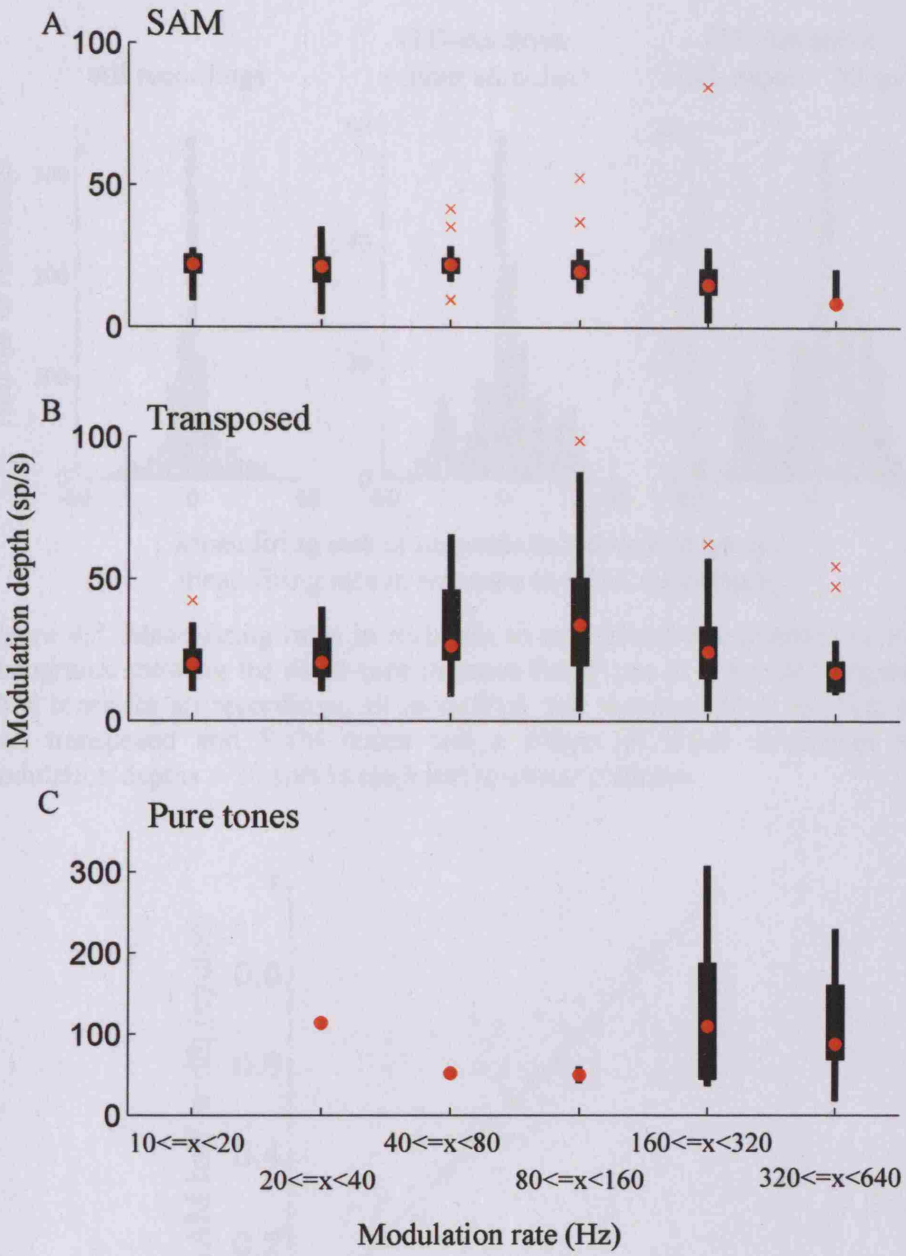

Figure 4.2. Box and whisker plots showing the distribution of modulation depths of all ITD-sensitive recordings, grouped according to the frequency of the stimulus. Symbols as in Figure 4.1.

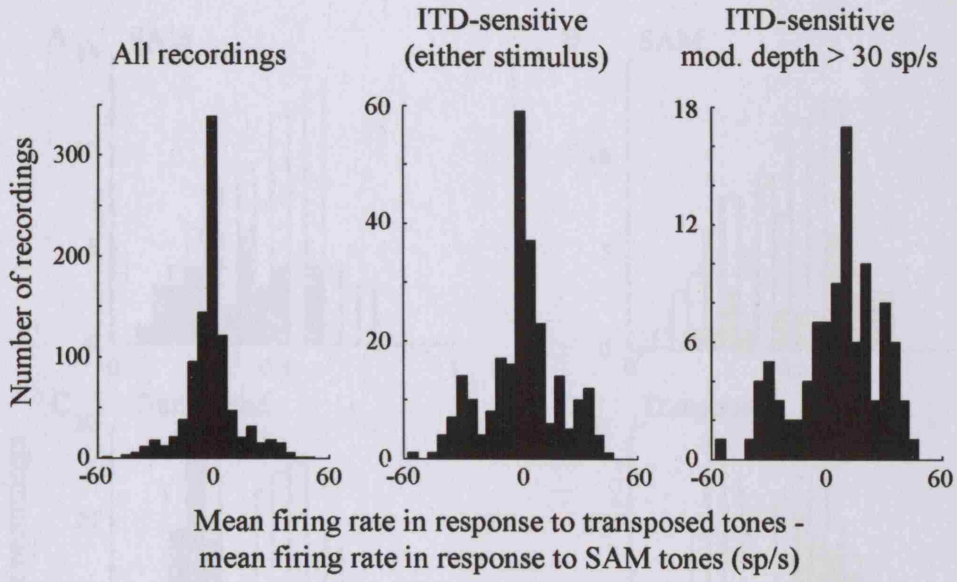

Figure 4.3. Mean firing rates in response to transposed as compared to SAM tones. Histograms showing the difference in mean firing rate in response to transposed and SAM tones for all recordings, all recordings that were sensitive to ITDs in either or both transposed and SAM tones and a subset of those recordings which had modulation depths > 30 sp/s in response to either stimulus.

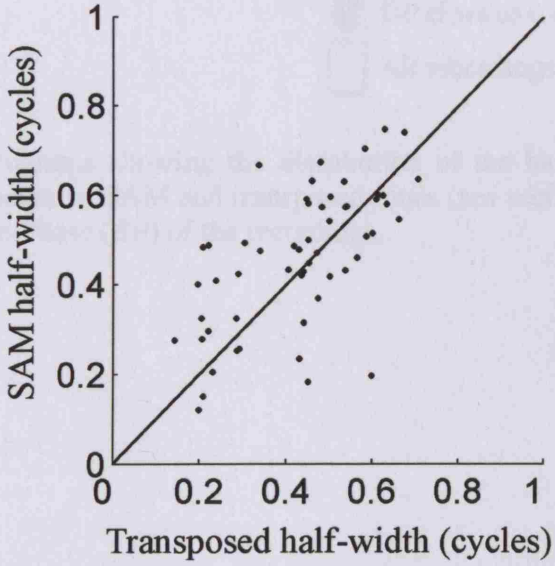

Figure 4.4. Half-widths were similar in response to SAM and transposed tones. Only ITD-tuning functions that were sensitive to ITDs in response to both SAM and transposed tones are shown. The diagonal line indicates equality between the halfwidths of the ITD functions in response to the transposed and SAM tones.

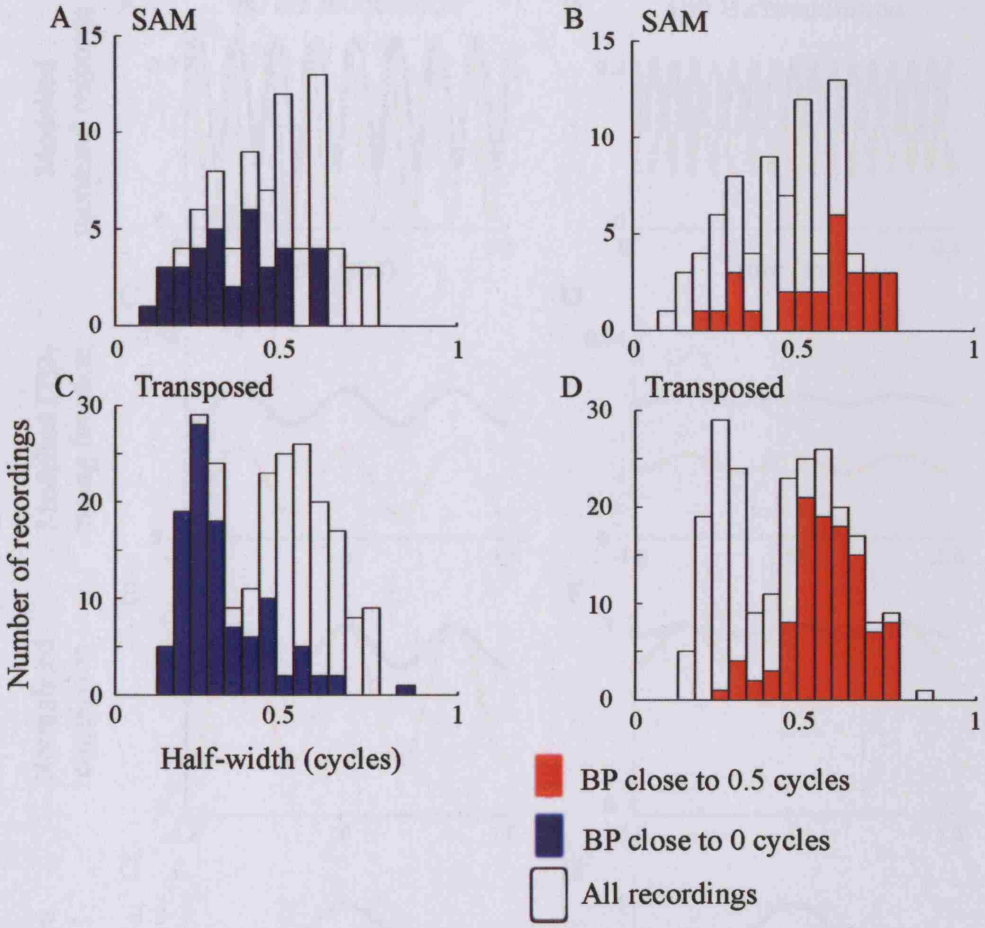

Figure 4.5. Histograms showing the distribution of the half-widths of ITD-tuning functions in response to SAM and transposed tones (see section 4.1). The Figure key indicates the Best Phase (BP) of the recordings.

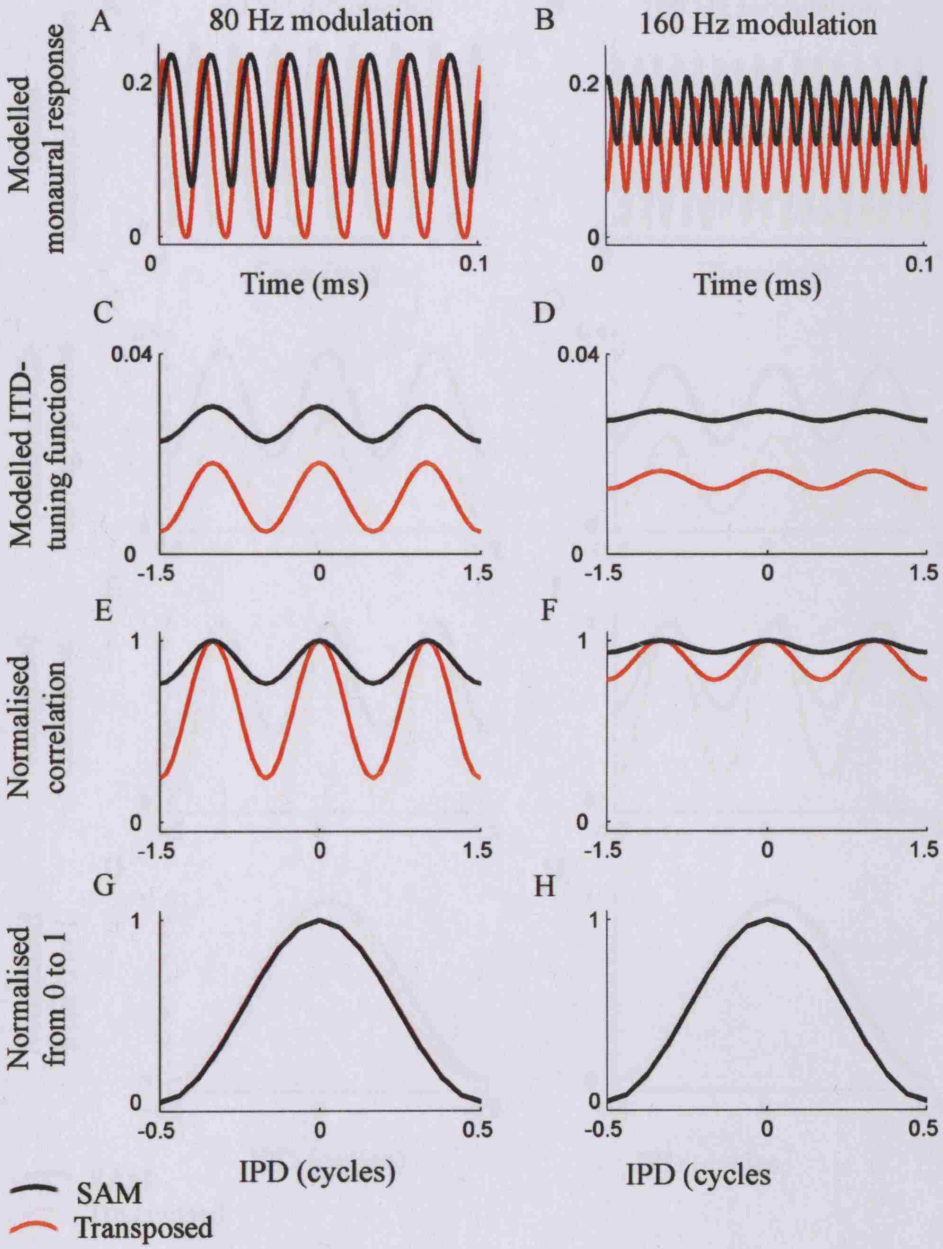

Figure 4.6. A) The result of modelling peripheral processing of a transposed (red) and SAM (black) tone centred at 4 kHz and with 80 Hz (left) or 160 Hz (right) modulation rate. Peripheral processing was modelled with 1. a gamma-tone filter centred at 4 kHz and 2. by half-wave, square-law rectification, envelope compression and low-pass filtering. Sensitivity to ITDs was modelled by then finding the cross-product (C & D), the normalised correlation (E & F) or by normalising the cross-product (or normalised correlation) from 0 to 1 (G & H).

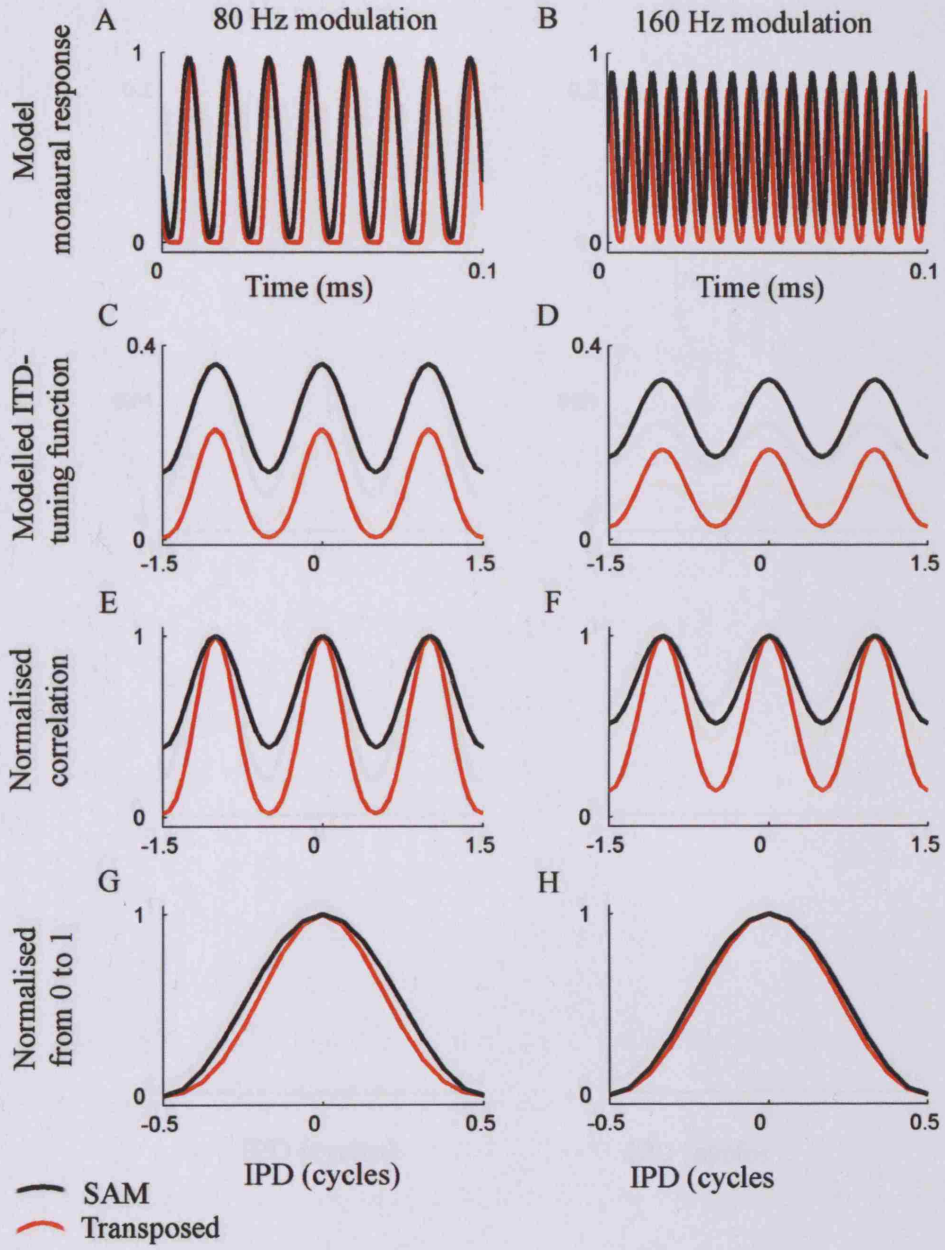

Figure 4.7. A) The result of modelling peripheral processing of a transposed (red) and SAM (black) tone centred at 4 kHz and with 80 Hz (left) or 160 Hz (right) modulation rate. Peripheral processing was modelled with 1. a gamma-tone filter centred at 4 kHz and 2. by taking the Hilbert envelope of the resulting signal. Sensitivity to ITDs was modelled by then finding the cross-product (C & D), the normalised correlation (E & F) or by normalising the cross-product (or normalised correlation) from 0 to 1 (G & H).

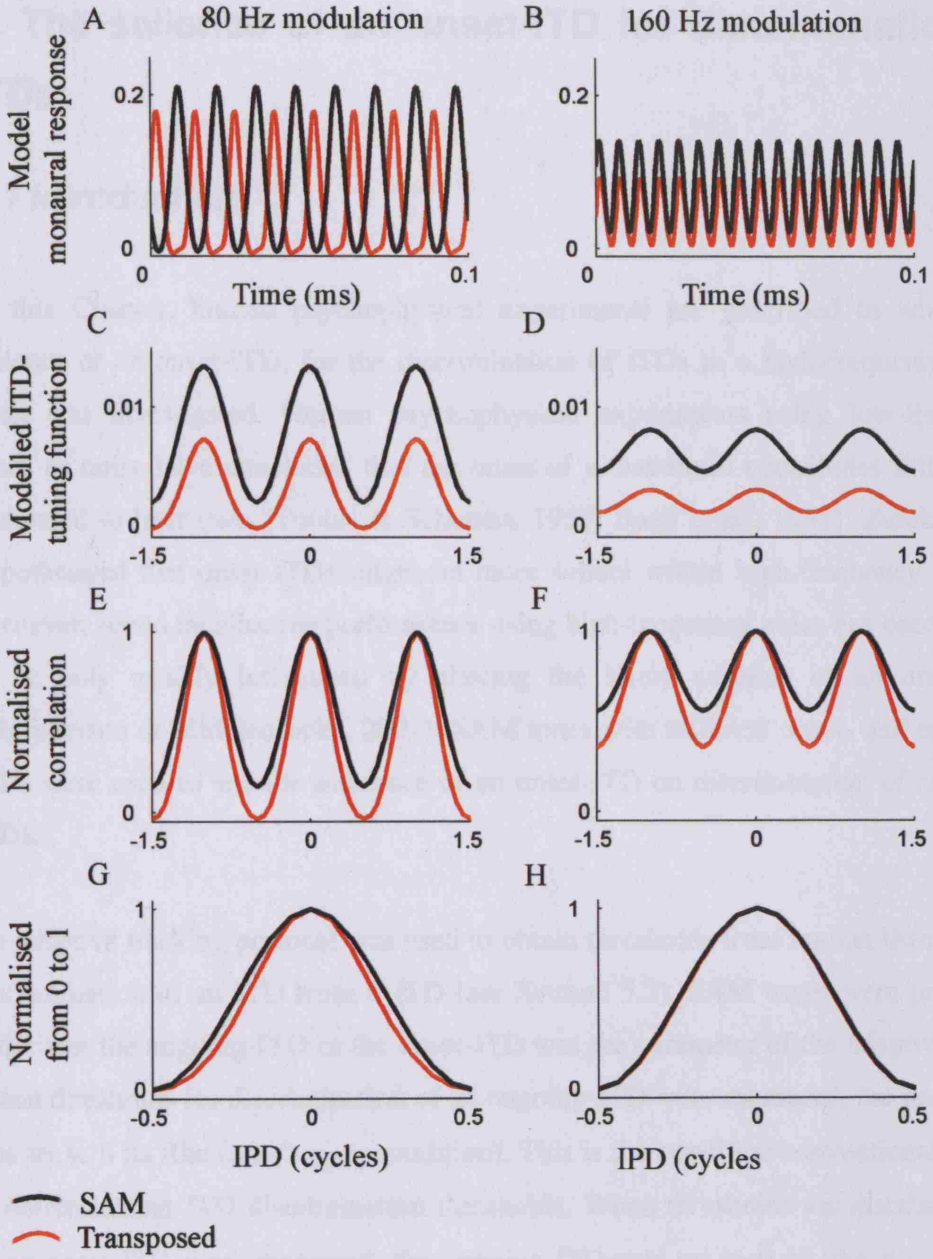

Figure 4.8. A) The result of modelling peripheral processing of a transposed (red) and SAM (black) tone centred at 4 kHz and with 80 Hz (left) or 160 Hz (right) modulation rate. Peripheral processing was modelled with 1. a gamma-tone filter centred at 4 kHz and 2. by half-wave, square-law rectification and low-pass filtering. Envelope compression was not modelled. Sensitivity to ITDs was modelled by then finding the cross-product (C & D), the normalised correlation (E & F) or by normalising the cross-product (or normalised correlation) from 0 to 1 (G & H).

5. The salience of an onset-ITD for discrimination of ITDs

5.1 Introduction

In this Chapter, human psychophysical experiments are described in which the salience of an onset-ITD, for the discrimination of ITDs in a high-frequency SAM tone, was investigated. Human psychophysical experiments using low-frequency tones or noise have concluded that the onset of a waveform contributes little when compared to later cues (Tobias & Schubert, 1959; Buell *et al.*, 1991). Zurek (1993) hypothesised that onset-ITDs might be more salient within high-frequency sounds. However, sound localisation performance using high-frequency noise has been shown to be only weakly influenced by altering the likely salience of an onset-ITD (Macpherson & Middlebrooks, 2002). SAM tones with different onset- and ongoing-ITDs were used to test the influence of an onset-ITD on discrimination of ongoing-ITDs.

An adaptive tracking protocol was used to obtain thresholds from human listeners for discrimination of an ITD from 0 ITD (see Section 5.2). SAM tones were presented and either the ongoing-ITD or the onset-ITD was the parameter of the adaptive track. When thresholds for discrimination of an ongoing-ITD were measured, the onset-ITD was set to 0 μs (the $0_{onset}/\tau_{ongoing}$ condition). This is the condition conventionally used to obtain human ITD-discrimination thresholds. When thresholds for discrimination of an onset-ITD were measured, the ongoing-ITD was set to 0 μs (the $\tau_{onset}/0_{ongoing}$ condition).

An onset-ITD might enhance or impair discrimination of an ongoing-ITD, depending on whether it is of the same sign as the ongoing-ITD. To investigate this, thresholds for discrimination of an ongoing-ITD were measured in two further conditions. In the $\tau_{onset}/\tau_{ongoing}$ condition, the onset-ITD and the ongoing-ITD were equal and the delay

in the waveform could also be referred to as a whole waveform delay. If onset-ITDs are a salient cue for discrimination of ITDs, thresholds are predicted to be smaller in the $\tau_{onset}/\tau_{ongoing}$ condition than in the $0_{onset}/\tau_{ongoing}$ condition. In the final condition ($r_{onset}/\tau_{ongoing}$), the onset-ITD, on each stimulus presentation, was selected randomly from a uniform distribution of values between ± 1 ms (in steps of 20 μs). In this case, if onset-ITDs are completely dominant over the ongoing-ITDs, for ITD-discrimination, it would not be possible to obtain a threshold for discrimination of an ongoing-ITD. In contrast, if onset-ITDs have no influence on discrimination of an ongoing-ITD, thresholds for discrimination of ongoing-ITDs would be equal irrespective of the onset-ITD. The four different stimulus conditions are summarised in Table 3.

Examples of the four conditions are shown diagrammatically in Figure 5.1A for a 50 ms SAM tone modulated at a rate of 125 Hz. In these examples, the ITD to be detected by the listeners is 1 ms. For three of the conditions, where the parameter of the adaptive track was the ongoing-ITD ($\tau_{ongoing}$), the 1 ms ongoing-ITD is easily observed; the right (green) waveform lags the left (blue) waveform by 1 ms. The ITD is present in both the envelope and the carrier of the waveform. In contrast, in the condition in which the onset-ITD is the parameter of the adaptive track, the ongoing-ITD is 0 μs ($\tau_{onset}/0_{ongoing}$). The monaural phase of both the carrier and the modulator (the phase of the carrier and envelope at onset and their phase in relation to each other) were randomly assigned for each stimulus presentation. As a result of this, the peak amplitudes in the ongoing portions of the stimuli occur at different time-points.

To examine the onsets of the waveforms in more detail the first 2 ms are shown in Figure 5.1B. Note that, although the offsets are not shown in detail, the offset ITDs followed the onset-ITDs. In condition $0_{onset}/\tau_{ongoing}$, the waveforms both start at 0 ms. The difference in the left and right waveforms at the onset is solely due to the ongoing-ITD in the waveform. In condition $\tau_{onset}/\tau_{ongoing}$, the waveforms have an onset-ITD equal to the ongoing-ITD. This is a whole waveform delay. In condition

$r_{onset}/\tau_{ongoing}$, the onset-ITD was "roved" by selecting an onset-ITD randomly from a uniform distribution of possible onset-ITDs between -1 (right leading) and +1 (left leading) ms in 20 μs steps. In Figure 5.1B, the roved onset-ITD was -0.48 μs. This value was chosen to illustrate that the onset of the right waveform can occur before the onset of the left waveform, but this not always the case. The rove was chosen to cover a wide range (± 1 ms) to maximize the likelihood of observing a disruption in the detection of the ongoing-ITD.

Thresholds were obtained using SAM tones of either 50 ms or 300 ms duration and with f_m of either 125 Hz or 250 Hz. It was predicted that the onset-ITD would be most salient in conditions where thresholds for discrimination of ongoing-ITDs have previously been shown to be largest i.e. for the shortest duration stimulus and for f_m = 250 Hz (e.g. Nuetzel & Hafter, 1976; Bernstein & Trahiotis, 2002).

5.2. Adaptive tracking procedure

Thresholds for detection of an ongoing-ITD ($\tau_{ongoing}$) of an onset-ITD ($\tau_{onset}/0_{ongoing}$) were determined using a modified two-cue, two alternative, forced choice, adaptive task (Trahiotis *et al.*, 1990). For each adaptive track, the stimuli were presented with the same onset condition, duration and f_m. Each trial consisted of a 200-ms warning interval, an 800-ms pause and four observation intervals in which a stimulus was presented, separated by 400 ms. The durations of the observation intervals were equal to the stimulus duration. Each interval was marked visually by a computer monitor. The listener had to detect the presence of a positive ITD (left-ear leading) that was presented with equal *a priori* probability in either the second or third intervals. In the remaining observation intervals, the stimuli had 0-μs ITD, except for the onset-ITD in the r_{onset} condition which was always randomly sampling from a uniform distribution. Feedback was provided for 800 ms after the listener responded. A further 800 ms elapsed before the next trial.

The values of the onset-ITDs were necessarily limited to multiples of the sample period i.e. 20 μs for a 50 kHz sampling rate and 10 μs for a 100 kHz sampling rate. Therefore, the ongoing-ITDs were also limited to multiples of the sample period. This produced a small quantization error (< 10%) at some ITDs. The ITD (either onset or ongoing) for which a threshold was measured on any given trial, was determined adaptively in order to estimate 70.7 % correct (Levitt, 1971). Step sizes in the adaptive track were selected from a geometric scale and thresholds were estimated as the geometric mean from the final 8, out of 16, reversals. Any thresholds 2 ms or above were considered as an inability to complete the task. These thresholds were set at 2 ms for statistical analysis. A multi-factor, unbalanced ANOVA was used to assess whether there was any effect of subject, f_m, duration or onset-type on the threshold ITDs obtained.

Preliminary data from two subjects are presented in this chapter. These experiments were conducted at the University of Connecticut Health centre, under the supervision of Professor Leslie Bernstein (see Acknowledgements). Two normal-hearing female adults, including the author, served as subjects. The author had some experience with binaural detection tasks. The other listener was very experienced (> 10 years) in a range of binaural psychophysical tasks. Three consecutive thresholds were obtained from each listener, for each of the 3 $\tau_{ongoing}$ conditions. Thresholds for f_m = 125 Hz were obtained before f_m = 250 Hz. Within each f_m, stimulus conditions were chosen in random order. Then, three more thresholds were obtained revisiting the same stimulus conditions in reverse order. Thresholds for the $\tau_{onset}/0_{ongoing}$ condition were obtained in the same manner, after the thresholds for the $\tau_{ongoing}$ conditions. There were deviations from this procedure when extra thresholds (more than 3) were obtained. All thresholds were included in the data analysis.

5.3. Results

Mean thresholds for discrimination of an ITD, from 0-μs ITD, are shown for subject X in Figure 5.2A and subject Y in Figure 5.2B. Thresholds are shown for discrimination of ongoing-ITDs (when the onset-ITD was 0 μs; $0_{onset}/\tau_{ongoing}$) and for discrimination of onset-ITDs (when the ongoing-ITD was 0 μs; $\tau_{onset}/0_{ongoing}$). The blue and black symbols mark thresholds obtained with SAM tones with a f_m of 250 Hz and 125 Hz, respectively. The stimulus durations and stimulus conditions are marked below each Figure (5.1A & B). The data in Figure 5.2 demonstrate that it is possible to discriminate between onset-ITDs as well as ongoing-ITDs, within high-frequency SAM tones. The mean thresholds obtained from subject X were higher for discrimination of an onset-ITD ($\tau_{onset}/0_{ongoing}$) than an ongoing-ITD ($0_{onset}/\tau_{ongoing}$). This was also the case for subject Y when the stimulus duration was 300 ms, or when the stimulus duration was 50 ms and the f_m was 125 Hz. However, when the stimulus duration was 50 ms and the f_m was 250 Hz, the mean thresholds from subject Y were higher for discrimination of an ongoing-ITD than an onset-ITD.

Mean thresholds for discrimination of ongoing-ITDs ($0_{onset}/\tau_{ongoing}$) were lowest when the f_m was 125 Hz as compared to 250 Hz, and when the stimulus duration was 300 ms as compared to 50 ms, consistent with previous data (e.g. Nuetzel & Hafter, 1976; Bernstein & Trahiotis, 2002). In contrast, mean thresholds for discrimination of onset-ITDs ($\tau_{onset}/0_{ongoing}$) were lowest when the f_m was 250 Hz as compared to 125 Hz. Also in contrast to thresholds for ongoing-ITDs, mean thresholds for discrimination of onset-ITDs were also lowest when the duration was 50 ms as compared to 300 ms for subject X. However, the duration of the stimulus had little effect on subject Y's thresholds for discrimination of onset-ITDs. The results shown in Figure 5.2 are consistent with the hypothesis that onset-ITDs are most salient when ongoing-ITDs are weak.

Mean thresholds for discrimination of ongoing-ITDs, when the onset-ITD was 0, are shown again in Figure 5.3. In addition, mean thresholds are shown for detection of ongoing-ITDs in two further conditions: $\tau_{onset}/\tau_{ongoing}$ and $r_{onset}/\tau_{ongoing}$. Thresholds in response to SAM tones with f_m = 125 Hz (black squares) were lower than thresholds in response to SAM tones with f_m = 250 Hz (blue circles), consistent with previous reports e.g. Bernstein & Trahiotis (2002). Also consistent with previous data, thresholds were smaller in response to a 300-ms stimulus compared to a 50-ms stimulus (Tobias & Zerlin, 1959).

An unbalanced, multi-factor ANOVA was performed to investigate the variance in the thresholds obtained. The result provided statistical evidence that all 4 of the experimental conditions (subject, duration, f_m and onset-ITD condition) affected thresholds for detection of ongoing ITDs ($p < 0.01$). The only interactions between these conditions that were not significant were the 4-way interaction and the interaction between subject, f_m and duration. Thresholds were elevated in response to stimuli with a roved onset-ITD compared to the other stimulus conditions (multiple comparison procedure with a bonferroni correction, $p < 0.01$). In the case of the 300-ms duration stimuli, it was still possible for both subjects to discriminate ongoing-ITDs at both 125 and 250 Hz f_m when the onset-ITD was roved. When the SAM tone was 50 ms in duration and the f_m was 250 Hz (i.e. the condition for which thresholds for an ongoing-ITD were largest) the roved onset-ITD was most disruptive. Subject Y was considered unable to do this task in 4/8 trials. (These thresholds were set to 2 ms for use in the statistical analysis.) However, when the f_m was 125 Hz and, particularly when the stimulus duration was 300 ms (i.e. conditions in which the smallest thresholds were obtained), the roved onset-ITD had little effect on ITD discrimination thresholds.

There was no significant difference between thresholds obtained in the $0_{onset}/\tau_{ongoing}$ compared to the $\tau_{onset}/\tau_{ongoing}$ conditions. There was a trend that thresholds were smaller in response to stimuli with a whole waveform delay ($\tau_{onset}/\tau_{ongoing}$). This was

especially noticeable when the f_m was 250 Hz. Detecting an ongoing-ITD when the f_m was 250 Hz was a more difficult task (i.e. thresholds were higher) compared to when the f_m was 125 Hz. That the onset-ITD was equal to the ongoing-ITD (as in the $\tau_{onset}/\tau_{ongoing}$ condition) may have provided a greater advantage for detection of an ITD when the ongoing-ITD was less informative.

5.4. Discussion

It was hypothesised that the onset-ITD of a SAM tone has little influence on the discrimination, from 0-μs ITD, of an ongoing-ITD. At least for the two subjects in this study, the onset-ITD had a weak influence on the discrimination of an ongoing-ITD, when the ongoing-ITD cues were strong. The onset-ITD was most salient if the ongoing-ITD cue was weak.

It was found that the onset-ITD of a SAM tone could influence the detection of an ongoing-ITD but only when the onset-ITD was roved by ± 1 ms; there was no significant difference between thresholds for detection of ongoing-ITDs when the onset-ITD was 0 μs or when the onset-ITD was equal to the ongoing-ITD. The rove in ITD was large compared to thresholds for ongoing-ITDs, potentially producing an onset-ITD that was 10 times greater than the threshold for an ongoing-ITD, and of opposite sign. However, this roved onset-ITD had little effect on the detection of an ongoing-ITD when the ongoing-ITD cue was strong (f_m = 125 Hz). It was most disruptive to the detection of an ongoing-ITD when the ongoing-ITD cue was relatively weak (i.e. at f_m = 250 Hz and duration = 50 ms). Consistent with the onset-ITD being a relatively weak cue for detection of an ITD, thresholds for detecting an onset-ITD were generally higher than those for detecting an ongoing-ITD, and were smallest when the ongoing-ITD was relatively weak (f_m = 250 Hz; Figure 5.2).

It was interesting to note the thresholds obtained from subject X were less influenced by the onset-ITD than from subject Y and, consistent with this, subject X had the

highest thresholds for detection of an onset-ITD. The results in this Section were only based on two listeners with different degrees of listening experience. These data should, therefore, be considered as preliminary.

5.4.1. Comparison with previous psychophysics

The influence of an onset-ITD on the salience of ITDs for localisation of a high-pass filtered noise (4-16 kHz) was investigated by MacPherson & Middlebrooks (2002) using virtual acoustic space (VAS; see Section II.2.3). When the potential salience of the onset-ITD was manipulated by altering the rise time, the weight that listeners attributed to ITDs was not affected. This suggests that, as found in this study, sensitivity to ongoing-ITDs is dominant over onset-ITDs within high-frequency stimuli.

Within low-frequency tones and noise, it has also been concluded that an onset-ITD has little influence on detection of ongoing-ITDs (Tobias & Schubert, 1959; Buell *et al.*, 1991). Buell *et al.* (1991) suggested that onset-ITDs, at low-frequencies, are effective only when the ongoing-ITD information is ambiguous. They showed that an onset-ITD can enable listeners to disambiguate an ongoing-ITD equal to 0.5 cycles of a low-frequency tone stimulus. At 0.5 cycles, an ITD can be lateralised towards either the left or right ear with equal probability. The percept can switch spontaneously between the two positions. An onset-ITD favouring the left ear can stabilise the percept towards that ear (Buell *et al.*, 1991).

Zurek (1993) tried to resolve the apparent inconsistency between data which suggest that an onset-ITD is less salient than the ongoing-ITD (Tobias & Schubert, 1959; Buell *et al.*, 1991), and studies pertaining to the precedence effect which suggest that the first arriving sound waveform dominates localisation and lateralisation (Litovsky *et al.*, 1999). Zurek (1993) noted that in studies where sensitivity to onset-ITDs was shown to be dominant over ongoing-ITDs, the onset-ITDs were more than a gating

delay. For example, in experiments conducted by Houtgast & Aoki (1994) and Zurek (1980) the stimulus, in an initial portion, contained conflicting spatial cues to the remainder of the stimulus. Zurek (1993) suggested that carrier ITDs in the initial portion of a waveform are also ongoing-ITDs and can dominate the ongoing-ITD in the rest of the waveform, whereas an onset- (or gating) ITD cannot. The shortest duration of an "onset-ITD", in studies where onsets are shown to dominate, refers to the duration of a single click. A click causes ringing on the basilar membrane and a response longer than the stimulus in ANFs and therefore can be considered more than a gating delay (Lin & Guinan, 2000). Freyman *et al.* (1997) present data that does not support Zurek's hypothesis. The conditions under which ITDs in the initial portion of a sound waveform are most salient were investigated using trains of short pulses, noise bursts or inharmonic complexes. Freyman *et al.* (1997) show that, even when the first pulse contains ITDs that are strong cues for lateralisation, the lateralisation of the sound can be dominated by ITDs in the ongoing waveform. When the ongoing-ITDs are ambiguous, the ITDs in the initial portion of the waveform were more likely to dominate.

Zurek (1993) asked, "what happens to onset effects (i.e. dominance) if the stimulus is restricted to high-frequencies." He hypothesised that onset-ITDs might dominate sensitivity to ITDs within high-frequency sounds because there is no sensitivity to carrier-ITDs. The results presented in this Chapter show that onset-ITDs do not dominate sensitivity to ITDs at high-frequencies. Hafter and colleagues demonstrated that more ITD (or spatial) information is gained from the first click in a train of high-frequency clicks than subsequent clicks [for interclick intervals of < ~13 ms (Buell & Hafter, 1988; Hafter & Buell, 1990; Stecker & Hafter, 2002)]. However, these experiments do not demonstrate, as suggested by Zurek (1993), that the first high-frequency click dominates perception if the subsequent clicks contain conflicting cues. Previous studies have shown that it is possible to discriminate between ITDs in single high-frequency clicks with thresholds < 100 μs (Buell & Hafter, 1988; Hafter & Buell, 1990). The relatively high thresholds for detection of onset-ITDs obtained in

the present study may be attributed to the dominance of the ongoing-ITD which was 0 μs.

Thresholds for detection of an ongoing-ITD were lowest in the present study when the stimulus duration was longest, consistent with the findings of other authors, for both high- and low-frequency stimuli (Tobias & Zerlin, 1959; Houtgast & Plomp, 1968; Hafter *et al.*, 1979; Buell & Hafter, 1988). The roved onset was most disruptive of listeners' ability to detect an ongoing-ITD when the stimulus duration was shortest (50 ms). This is consistent with the results of Tobias & Schubert (1959) who asked listeners to centre the lateralisation of broadband noise by controlling the value of the ongoing-ITD. They found that an onset-ITD needed to be offset by a larger ongoing-ITD for a short duration noise than for a noise of longer duration. When the duration of the noise was 300 ms or more, an onset-ITD up to 400 ms was almost completely dominated by an ongoing-ITD of 0 μs.

Hafter *et al.* (1979) measured thresholds for detection of an ongoing-ITD in low-frequency tones as a function of duration, in the presence and absence of a noise which masked the onset and offset transients. Removal of the onset-ITD raised thresholds for detection of an ITD for short duration signals (below ~100 ms). This could be interpreted to indicate that onset-ITDs are most salient for short duration stimuli although the authors suggest that the elevated thresholds may be explained by forward and backward masking of the tone by the noise.

Thresholds for detection of an onset-ITD ($\tau_{onset}/0_{ongoing}$) within the current experiment did not appear to depend greatly on the duration of the SAM tone although subject X showed a trend towards lower thresholds with 50-ms stimuli. Perrott & Baars (1974) report that thresholds for detecting onset-ITDs from 0-μs ITD were lowest (indicating greatest sensitivity to the onset-ITD) for the shortest signal duration (1 ms) and increased as the signal duration increased (to 1000 ms). This was interpreted to indicate that the ongoing-ITD was increasingly salient, compared to the onset-ITD, as

signal duration increases. That the salience of an onset-ITD does depend on the duration of the ongoing stimulus was supported by the observation that the roved onset ($r_{onset}/\tau_{ongoing}$) was most disruptive to the detection of an ongoing-ITD when the stimulus duration was 50 ms. These observations are also consistent with the notion that onset-ITDs are most salient when ongoing-ITD cues are weak.

Perrott & Baars (1974), in their study measuring thresholds for detection of onset and offset ITDs, noted that, for stimuli of 100 and 1000 ms, these cues were reported by the listeners to be heard as a monaural transient. This was not reported by the subjects in the present study.

5.4.2. Discussion with reference to neural responses

The finding that an onset-ITD does not have a dramatic influence on sensitivity to ongoing-ITDs is consistent with the physiological observation that onset-ITDs have little influence on ITD-tuning functions calculated from the mean firing rates of neurons in the IC of the guinea pig (Griffin *et al.*, 2005, Chapter 3 and Section I.1.5). In the experiments presented in Chapter 3, sensitivity to ITDs was assessed with a whole waveform delay. Neural thresholds for detection of ITDs within high-frequency signals were generally similar whether the response to the onset of the stimuli was included in the calculation of neural firing rates or not (see Section I.1.5). Firing rates were obtained at + 0.5 cycles and – 0.5 cycles IPD. Waveforms with these two IPDs have the same ongoing-ITD but different onset- and offset- ITDs. Firing rates at ± 0.5 cycles IPD were similar, suggesting that the onset-ITD did not influence the firing rate of the neurons studied (see Section I.1.3). Kuwada and Yin (1983) found, in their study investigating the sensitivity of low-frequency IC neurons to ITDs, that most neurons (93%) are sensitive to ongoing but not to onset-ITDs.

Heil and Neubauer (2001, 2003) have investigated the factors that influence first spike latency in response to auditory stimuli. They found that first-spike latency is

related to the integral of the sound pressure waveform. The phase of the envelope (and the carrier), the intensity and the onset ramp of a SAM tone will affect the integral of the sound pressure waveform and timing of the first sound-evoked spike. The carrier and envelope phase were both randomised between trials, minimising the effect of these parameters on the sound pressure integral. However, the dependence of the results presented here on the intensity of the stimulus and the duration of the onset (and offset) ramp is unknown. The stimuli in the experiments presented in this Chapter were ramped on with a 10 ms $\cos^2$ ramp. The onset-ITDs might have been more salient if the stimuli were ramped on more quickly (Gardner, 1968; Abel & Kunov, 1983; Rakerd & Hartmann, 1986; Litovsky *et al.*, 1999). However, a fast onset creates a broad spectrum that could provide ITDs across a broad range of frequencies, providing an advantage in ITD discrimination due to across-frequency integration or the use of ITDs in low-frequency sound. Further experiments might include variation in the duration of the onset ramp to investigate whether this was critical to the results presented here.

5.4.3. Physiological relevance

Natural high-frequency sounds contain ITDs, IIDs, and spectral notches, which can be used for sound localisation. The current study only investigated sensitivity to ITDs. This caveat aside, the results suggest that, in the case when ongoing-ITD information is confusing (such as in a highly reverberant environment), onset-ITDs cues are useful for localising high-frequency sounds, otherwise ongoing-ITDs dominate onset-ITDs.

5.5 Figures and Tables for Chapter 5

Stimulus Condition	*Onset (or gating) ITD*	*Ongoing ITD*
$0_{onset}/\tau_{ongoing}$	Zero	Parameter of adaptive track
$\tau_{onset}/0_{ongoing}$	Parameter of adaptive track	Zero
$\tau_{onset}/\tau_{ongoing}$	Equal to ongoing	Parameter of adaptive track
$r_{onset}/\tau_{ongoing}$	Roved by +/- 1 ms	Parameter of adaptive track

Table 3. A summary of the properties of the onset- and ongoing ITDs of the SAM tones used as stimuli for the data presented in chapter 5.

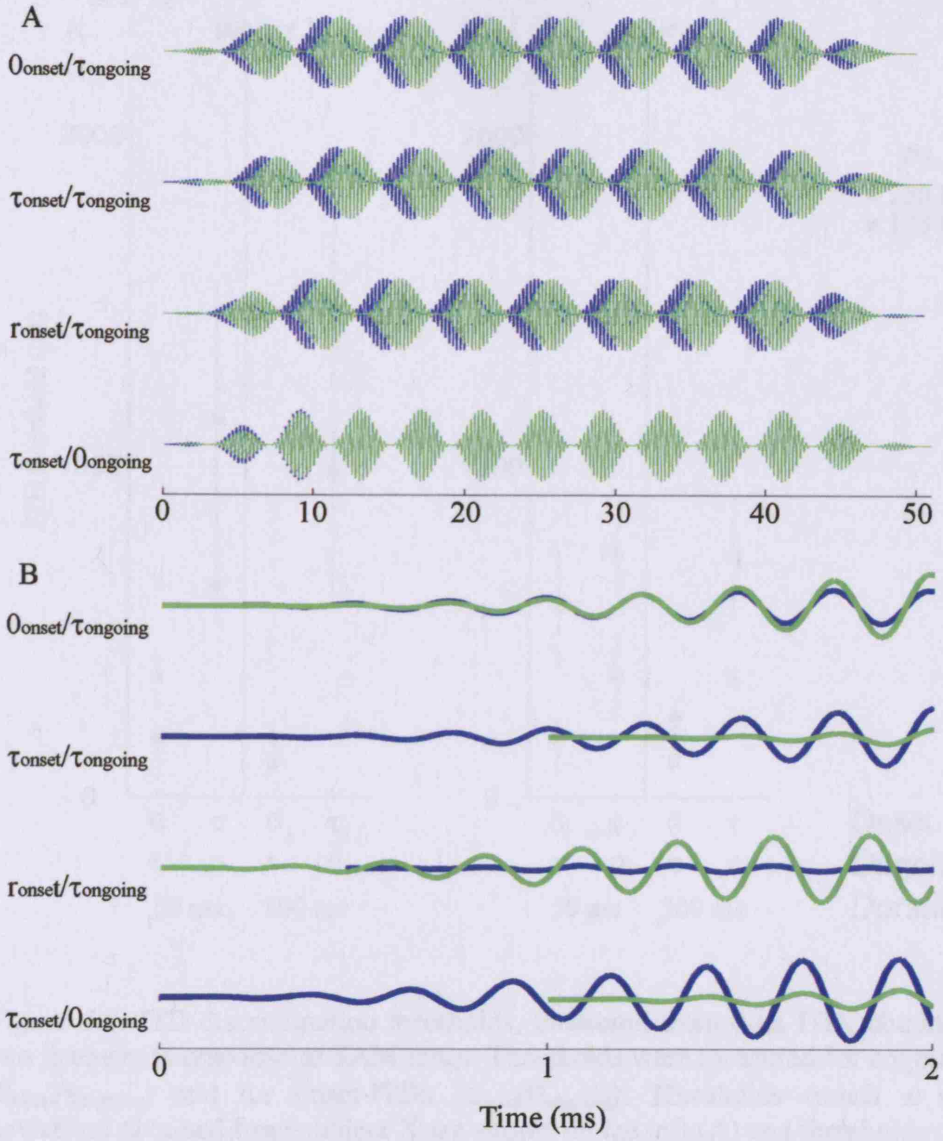

Figure 5.1. SAM tones with different onset-ITD conditions. The legends are expanded in Table 3. A) Four 50 ms SAM tones centred at 4 kHz and modulated by 125 Hz illustrating the onset- and ongoing-ITD conditions used in the study described in Chapter 5. B) The first 2 ms of the waveforms in A. The scales of the ordinates in B are not necessarily equal. The blue and green waveforms represent stimuli sent to the left and right headphones, respectively.

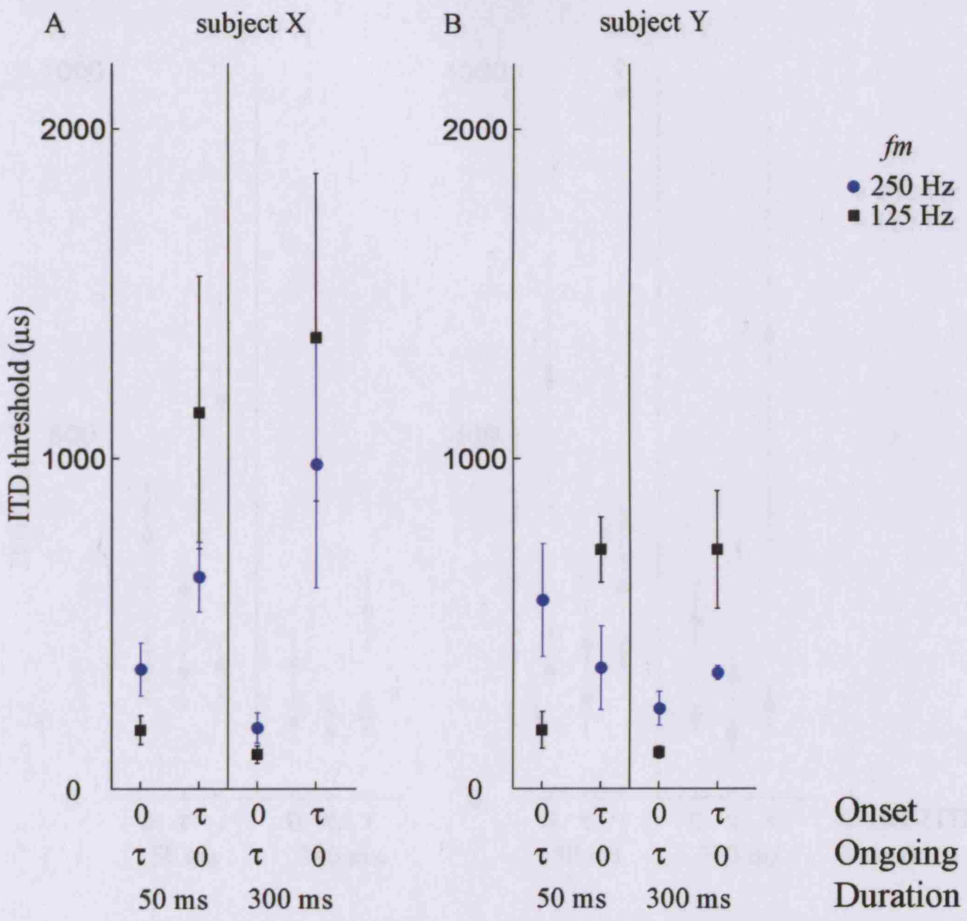

Figure 5.2. ITD discrimination thresholds, measured from 0-μs ITD, obtained from two listeners in response to SAM tones. Thresholds were measured for ongoing-ITDs ($0_{onset}/\tau_{ongoing}$) and for onset-ITDs ($\tau_{onset}/0_{ongoing}$). Thresholds (mean ± standard deviation) obtained from subject X are shown on the left (A) and thresholds obtained from subject Y are shown on the right (B). The thresholds (mean ± standard deviation) from each subject are divided according to the duration of the stimulus (50 ms: left, 300 ms: right). Blue circles represent mean thresholds for SAM tones with 250 Hz f_m whereas black squares represent mean thresholds for transposed tones with 125 Hz f_m.

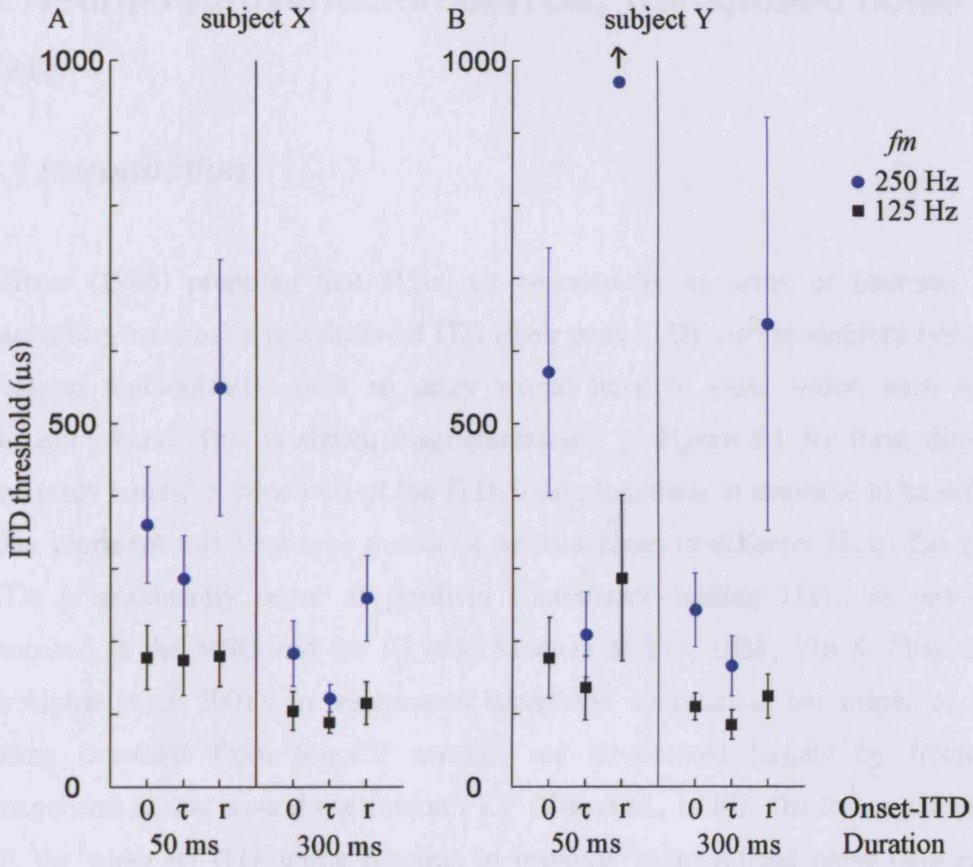

Figure 5.3. ITD discrimination thresholds, measured from 0-µs ITD, obtained from two listeners in response to SAM tones. Thresholds were measured for ongoing-ITDs and for three different onset-ITD conditions: $0_{onset}/\tau_{ongoing}$, $\tau_{onset}/\tau_{ongoing}$, and $r_{onset}/\tau_{ongoing}$. Thresholds (mean ± standard deviation) obtained from subject X are shown on the left (A) and thresholds obtained from subject Y are shown on the right (B). The thresholds (mean ± standard deviation) from each subject are divided according to the duration of the stimulus (50 ms: left, 300 ms: right). Blue circles represent mean thresholds for SAM tones with 250 Hz f_m whereas black squares represent mean thresholds for transposed tones with 125 Hz f_m.

6. Neural representation of ITDs; Transposed noise data

6.1 Introduction

Jeffress (1948) proposed that ITDs are encoded by an array of neurons, each responding maximally to a different ITD (their peak-ITD). As the auditory system is arranged tonotopically, such an array would have to exist within each sound frequency band. This is shown diagrammatically in Figure 6.1 for three different frequency bands, in which *all* of the ITD-tuning functions in response to broadband noise represent a Jeffress-type matrix of neurons tuned to different ITDs. The peak-ITDs predominantly occur at positive, contralateral-leading ITDs, as has been described in the MSO and the IC (e.g. Kuwada & Yin, 1983; Yin & Chan 1990, McAlpine *et al.* 2001). In response to broadband stimulation, the shapes of ITD-tuning functions from low-CF neurons are determined largely by frequency components at and around the neuron's CF (Yin *et al.*, 1986). The lower a neuron's CF, the wider its ITD-tuning function in response to broadband noise (Yin *et al.*, 1986). More recent evidence suggests that the peak-ITDs of ITD-tuning functions are not distributed equally across CFs but are "CF-dependent" (McAlpine *et al.*, 1996; McAlpine *et al.*, 2001; Brand *et al.*, 2002; Hancock & Delgutte, 2004). Neurons with the highest CFs tend to respond maximally to the shortest ITDs, as represented by the *red* ITD-tuning functions in Figure 6.1. Peak-ITDs are therefore proportional to the width of ITD-tuning functions in response to broadband noise. This relationship has been suggested to maintain the greatest "dynamic range" of ITD-tuning functions over the range of ITDs that can be experienced (McAlpine *et al.* 2001).

In the current work, it was hypothesised that the frequency-filtering properties of neurons with high-frequency CFs, in terms of the envelope frequencies that dominate their ITD-tuning functions, determine their tuning for ITDs. If high-CF neurons demonstrate a frequency-dependent ITD-tuning similar to CF-dependent ITD-tuning

of low-CF neurons, the widest ITD-tuning functions from high-CF neurons will peak at the largest ITDs. It has previously been observed that the half-width of ITD-tuning functions in response to high-frequency SAM tones depends on the f_m of the SAM tone; larger f_m produce narrower ITD-tuning functions (Batra *et al.*, 1993). When a stimulus is narrowband, as for example in the case of a SAM tone, and assuming that a neuron is tuned to a constant ITD, the relationship between peak-ITD and half-width depends on the f_m of the stimulus (Batra *et al.*, 1993). Therefore, to investigate frequency-dependent ITD-tuning in high-CF neurons, a stimulus that was broadband for envelope frequencies was chosen so that ITD-tuning functions were dominated by neural tuning to envelope frequencies. A transposed noise was chosen instead of a Gaussian noise because it was previously found (see Chapters 3 and 4) that neural responses to ITDs were more strongly modulated in response to transposed tones compared to the more commonly used SAM tones. A token of the transposed noise stimulus used in the experiments described in Chapter 6 is shown in Figure 6.2. A low-pass (cut-off = 800 Hz), flat spectrum noise (Figure 6.2A) was half-wave rectified (Figure 6.2B) and multiplied by a high-frequency (4 kHz) carrier tone (Figure 6.2C). The spectrum of the resulting stimulus (Figure 6.2D) has a broad spread of energy centred on 4 kHz.

6.2. Data Analysis

Data were obtained from 65 high-CF neurons from a total of 24 animals. Eight of these animals were also recorded from in response to SAM and transposed tones.

ITD sensitivity

Transposed noise stimuli were either 200 or 1000 ms in duration and spikes were counted from 5 - 260 ms or from 5 - 1060 ms after stimulus presentation, respectively. The mean and standard deviation of firing rates at each stimulus condition was used in subsequent analysis. Modulation of neural firing rate, in

response to ITDs, was assessed by a non-parametric ANOVA (Friedman's test). Neurons were considered sensitive to ITDs if there was statistical evidence that their firing rates were not constant as a function of ITD ($p < 0.001$).

Peak-ITDs, trough-ITDs and half-widths and steepest slopes

The ITDs at the maximum and minimum of ITD-tuning functions were estimated by fitting a 5^{th} degree polynomial to 11 firing rates centred on the maximum (or minimum) firing rate and finding the ITD (to the nearest μs) associated with the maximum (or minimum) of the polynomial fit. Half-widths were calculated as described for ITD-tuning functions in response to transposed and SAM tones but were expressed in microseconds. The steepest slopes were calculated as described for ITD-tuning functions in response to transposed and SAM tones.

Composite functions

Composite functions are a model of ITD-tuning functions in response to noise stimuli and were calculated from neural firing rates in response to ITDs in transposed tones. The firing rates of the composite functions were calculated for ITDs between ± 4 ms, in 0.5 ms steps. At each ITD, the firing rate of a composite function for a single neuron was equal to the mean firing rate, at that ITD, in response to transposed tones with different f_m. However, ITD-tuning functions in response to transposed tones were not sampled in 0.5 ms steps at each f_m. Therefore, firing rates at ITDs that were not presented were obtained by linear interpolation. In addition, ITD-tuning functions in response to transposed tones were not obtained in response to ITDs as large as ± 4 ms at each f_m; the ITDs presented were uniformly sampled over a cycle of IPD. ITD-tuning functions in response to tones are periodic, so firing rates were replicated to encompass ± 4 ms. The calculation of the composite functions therefore involved a considerable degree of estimation.

Ideally, to compare composite functions derived from responses to transposed tones with ITD-tuning functions in response to transposed noise, responses to transposed tones should be sampled at uniformly spaced f_m, encompassing the same f_m present in the transposed noise. However, the sampling of f_m was different for each neuron and is noted below for each composite function (A-H) in Figure 6.2.1.4. The notation, "10:100:610" indicates transposed tones were presented at f_m from 10 to 610 Hz in 100 Hz steps. A) neuron 28104 – f_m = 10:100:510 Hz, B) neuron 27204 – f_m = 10:100:310 & 20:20:280 Hz, C) neuron 29005 – f_m = 10:100:610 Hz, D) neuron 28806 – f_m = 10:100:610 & 25:50:675 Hz, E) neuron 27202 – f_m = 10:100:410 & 20:20:140 Hz, F) neuron 29004 – f_m = 10:100:610 & 25:25:175 Hz, G) neuron 28803 – f_m = 10:100:410 & 50:25:400 Hz & H) 30212 – f_m = 10:100:610 Hz.

Data from low-frequency neurons

ITD-tuning functions obtained from low-CF neurons were previously published (McAlpine *et al.*, 2001). Data were analysed from 292 high-CF neurons from a total of 96 animals. Only ITD-tuning functions that were considered to be "peak-type" were analysed. This classification was based on the shape of the ITD-function if the side-peaks were small. If side-peaks were prominent, such that it was not clear whether an ITD-tuning function was dominated by a peak or trough in firing rate, responses to monaural stimulation was examined. Peak-type responses were characterised by excitation in response to sound at either ear. CFs and Peak-ITDs were estimated as described in McAlpine *et al.* (2001). Half-widths were estimated by estimating the ITD where firing rates were half way between the maximum and minimum firing rate (carried out by D. McAlpine).

6.3. Results

Action potentials from 65 neurons with high-CFs (median CF = 3.8 kHz, minimum = 2.2 kHz, maximum = 8.5 kHz) were recorded in response to transposed noise stimuli presented at a range of ITDs.

6.3.1. ITD-tuning in response to a high-frequency, broadband transposed noise

In the case of 63/65 high-CF neurons, the cut-off frequency for the low-pass, modulating noise was CF/5. This generated, when "transposed" to high-frequencies by multiplication with a tone equal to CF, a broadband transposed noise stimulus (Figure 6.2.1). From hereon, a transposed noise with f_c equal to CF and with the cut-off of the low-pass modulating noise equal to CF/5 will be referred to as "broadband transposed noise". Of 63 high-CF neurons, firing rates from 22 were considered sensitive to ITDs in response to broadband transposed noise (see Section 6.2). Note that this is likely to over-estimate the proportion of high-CF neurons that were sensitive to ITDs because, in some experiments, broadband transposed noise was only presented when neurons sensitive to ITDs within transposed tones were encountered.

Figure 6.3 shows ITD-tuning functions in response to broadband transposed noise from 8 neurons that were sensitive to ITDs. Firing rates were relatively weakly modulated as a function of ITD and never fell to 0 sp/s. The modulation depths (difference between the maximum and minimum firing rate) ranged between 3 and 34 sp/s with a median of 13 sp/s whereas the modulation depths recorded in response to transposed and SAM tones could be over 50 sp/s (Figure 4.2). The examples shown in Figure 6.3 include ITD-tuning functions that were amongst the most strongly modulated by ITDs and those that were very weakly modulated by ITDs.

ITD-tuning functions in response to pure, SAM and transposed *tones* are periodic. In response to *Gaussian noise*, ITD-tuning functions recorded from low-CF neurons are quasi-periodic; side-peaks and side-troughs are of smaller magnitude than a central peak and/or trough in firing rate (Geisler *et al.*, 1969; Yin *et al.*, 1986; Chan *et al.*, 1987; Yin *et al.*, 1987; McAlpine *et al.*, 2001; Joris *et al.*, 2005). This "damping" arises from a neuron's sensitivity to ITDs at different pure tone frequencies, which generate ITD-tuning functions with different periodicities (Yin *et al.*, 1986). ITD-tuning functions in response to transposed noise were generally highly damped. An ITD-tuning function that had clear side-peaks was only obtained from one neuron (Figure 6.3B).

ITD-tuning functions in response to broadband transposed noise were dominated by either a single peak (e.g. Figure 6.3A & B), a single trough (e.g. Figure 6.3C & D) or, in 2 cases from 2 different neurons, there was both a prominent peak and a prominent trough in firing rate (e.g. Figure 6.3F). If a neuron always fires maximally at the same ITD, irrespective of the frequency of the stimulus, the neuron's ITD-tuning function in response to broadband noise will be dominated by a peak at that ITD. This is predicted for a coincidence detector neuron which fires maximally in response to an ITD which is equal in magnitude but opposite in sign to the neuron's "internal delay" (see Figure 1.4 and Section 1.3). Similarly, if a neuron always fires minimally at the same ITD, irrespective of the frequency of the stimulus, its ITD-tuning in response to noise will be dominated by a trough. This is predicted for a coincidence detector neuron which detects the coincidence of an excitatory and inhibitory input and fires minimally at an ITD equal and opposite to its "internal delay" (see Section 1.3). However, if the ITD-tuning functions of an individual neuron in response to pure tones or single f_m do not peak or trough at a common ITD, ITD-tuning functions in response to noise may have a prominent peak and a prominent trough. This may occur if a coincidence detector neuron does not have a constant "internal delay" or if a neuron's response reflects convergent input from coincidence detector neurons with different properties (McAlpine *et al.*, 1998). ITD-tuning functions were classified as "peak-type" if the ITD at which the peak in firing rate occurred (peak-ITD) was

closer to 0-μs ITD than the ITD at which the trough occurred (trough-ITD). If the opposite was true, the ITD-tuning function was classified as "trough-type".

To examine the shapes of ITD-tuning functions in response to broadband transposed noise, smoothed ITD-tuning functions were normalised to their maximum and minimum firing rates. These are shown in Figure 6.4A. For neurons from which more than one ITD-tuning function was obtained (at different intensities), the data from the ITD-tuning function with the highest modulation in firing rate are used in this Chapter, unless stated otherwise. The maximum possible range of ITDs experienced by the guinea pig is indicated by the shaded region (Sterbing *et al.*, 2003). This maximum ITD, of ± 330 μs, occurred around 400-600 Hz and was smaller for high-frequency sounds (Sterbing *et al.*, 2003). Note that this is a wide estimate of the range of ITDs that a guinea pig might experience and has been explained by increases in the travel time of sound caused by the head and pinnae of the guinea pig. A spherical head model produces an estimate of the maximum ITD that a guinea pig can experience of ~150 μs. The maximum ITD measured from guinea pig HRTFs recorded in the laboratory in which this work was conducted, is 180 μs (N. J. Ingham – personal communication). Only 3 ITD-tuning functions were modulated by ≥ 60% of their modulation depth over ±330 μs ITD. ITD-tuning functions were even less modulated over a smaller range of ITDs.

Normalised ITD-tuning functions are shown again in Figures 6.4B & C according to whether they were classified as peak-type or trough-type, respectively. All peak-type ITD-tuning functions had peak-ITDs at positive, contralateral leading ITDs (sign test for median $\neq 0$ μs, $p < 0.001$). The steepest positive slopes of peak-type ITD-tuning functions tended to occur in a similar position, at ITDs just negative of 0 μs (median = -200 μs; sign test for median $\neq 0$ μs, $p < 0.01$), whereas the negative slopes were more widely distributed. This is reminiscent of the distribution of the steepest slopes of ITD-tuning functions (centred on 0-μs ITD) from low-CF neurons (McAlpine *et al.* 2001; Shackleton *et al.*, 2003).

The peak-ITDs and the half-widths of peak-type ITD-tuning functions in response to broadband transposed noise were estimated to investigate any "frequency-dependent" ITD-tuning of high-CF neurons (see Section 6.2). The CFs of low-CF neurons can be considered to "dominate" the width of ITD-tuning functions in response to a broadband stimulus; the half-width of ITD-tuning functions in response to broadband noise is a property of the frequency of sound to which a neuron responds and is positively correlated with CF (Yin *et al.*, 1986; Chan *et al.*, 1987). ITD-tuning functions obtained from high-CF neurons are dependent on the f_m at which sensitivity to ITDs occurs (Batra *et al.*, 1993; Joris & Yin, 1995). Therefore, a "dominant" f_m that can be considered to characterise the ITD-tuning functions in response to broadband transposed noise was estimated using their half-widths (see below).

The peak-ITDs of peak-type ITD-tuning functions obtained in response to broadband transposed noise are shown as a function of the half-width of the ITD-tuning functions in Figure 6.5A (black circles). Peak-type ITD-tuning functions with wider half-widths tended to peak at longer ITDs whereas peak-type ITD-tuning functions with narrower half-widths tended to peak at smaller ITDs (Figure 6.5A), similar to ITD-tuning functions from low-CF neurons (McAlpine *et al.*, 2001). A linear regression between the peak-ITD and the half-width of peak-type ITD-tuning functions was significant and accounted for 57% of the variance in peak-ITD (black line; $p = 0.0018$). The relationship between half-width and peak-ITD was robust to the method used to estimate these parameters (data not shown).

The trough-ITDs (ITD at the trough) of trough-type ITD-tuning functions are also shown as a function of half-width in Figure 6.5A (blue squares). The linear regression that was fit to these data was not significant ($p = 0.90$). Marquardt & McAlpine (2001) report that the highest peaks of trough-type ITD-tuning functions tend to occur at ~135° of the period of the CF. It was not possible to assess the peaks of the highly damped trough-type ITD-tuning functions in response to broadband transposed noise. However, neither the positive or negative slopes of trough-type ITD-tuning functions

were obviously clustered together (Figure 6.4C), suggesting that such a relationship does not exist in the small sample of trough-type ITD-tuning functions obtained.

Peak-ITDs (and trough-ITDs) are shown as a function of estimated dominant f_m in Figure 6.5C. Dominant f_m were estimated from the reciprocal of the half-width. The half-width of an ITD-tuning functions is a fraction of the period of the dominant frequency of an ITD-tuning function. Therefore, this reciprocal was multiplied by 0.25 for peak-type ITD-tuning functions and by 0.44 for trough-type ITD-tuning functions; 0.25 and 0.44 were the median half-widths of ITD-tuning functions obtained in response to transposed tones with a peak, or trough, in firing rate close to 0-μs ITD (see Figure 4.6). The range of estimated dominant f_ms, for the 22 ITD-tuning functions was 74 - 480 Hz. The peak-ITDs show a dependence of the estimated dominant f_m similar to that described in McAlpine *et al.* (2001) for low-CF neurons and can be described as "f_m-dependent".

6.3.2. Comparison with ITD-tuning functions from low-CF neurons in response to Gaussian noise

The peak-ITDs and half-widths of ITD-tuning functions in response to broadband Gaussian noise were examined from 292 IC neurons with CFs between 90 Hz and 1.8 kHz (median = 500 Hz). These data form part of a database of neural recordings from the same animal preparation used in the experiments described in here and were previously published in McAlpine *et al.* (2001). All the neural ITD-tuning functions from low-CF neurons were considered to be dominated by a peak in ITD (see Section 6.2). Figure 6.6 shows a comparison of these data with the ITD-tuning functions obtained in response to broadband transposed noise.

In Figure 6.6A (black circles), the peak-ITDs as a function of half-width in response to broadband transposed noise are re-plotted from Figure 6.6A. Alongside these data, the peak-ITDs of ITD-tuning functions in response to Gaussian noise are plotted as a

function of their half-widths (green circles). As described in Section 6.2.1, the peak-ITDs tended to be larger when the half-widths were larger, for ITD-tuning functions from both low- and high-CF neurons. The half-widths of ITD-tuning functions from high-CF neurons were never as narrow as many of the half-widths of ITD-tuning functions from low-CF neurons, reflecting the relatively low f_m at which sensitivity to ITD in the envelope of high-frequency sounds can occur.

Compared over a small range of half-widths, peak-ITDs tended to be lower for ITD-tuning functions recorded from high-CF as compared to low-CF neurons (Figure 6.6A). The same trend is observed when peak-ITDs are expressed as a proportion of the half-width (Figure 6.6B). Peak-ITDs expressed as a proportion of half-width were lower from high-CF as compared to low-CF neurons (Mann-Whitney U test, $p < 0.001$). This is consistent with the difference in the ITDs at which the steepest slopes occurred, which were also lower from high-CF neurons. The steepest slopes of ITD-tuning functions from low-CF neurons are distributed around 0-μs ITD (McAlpine *et al.* 2001) whereas the steepest positive slopes of ITD-tuning functions from high-CF neurons were distributed around ipsilateral leading (negative) ITDs (Figure 6.4B). In response to transposed noise neither the peak-ITDs nor the trough-ITDs were outside ± 0.5 times the half-width whereas peak-ITDs from low-CF neurons do not fall outside ± the half-width.

Figure 6.6C shows peak-ITD in response to broadband transposed noise as a function of estimated dominant f_m, re-plotted from Figure 6.2.1B (blue). Peak-ITDs of ITD-tuning functions from low-CF neurons are plotted in the same figure, but as a function of CF (green). The range of estimated dominant f_m was lower than the range of CFs, reflecting the lower range of f_m, as compared to the range of CFs, over which sensitivity to ITDs can occur. Peak-ITDs from high-CF neurons were generally lower than peak-ITDs from neurons with CFs within the same range as the estimated dominant f_ms. This partly reflects the lower peak-ITDs, as a proportion of half-width, from high-CF as compared to low-CF neurons (Figure 6.6B). In addition, this may

partly be explained if the dominant f_ms were under-estimated. The dominant f_ms were estimated using the median half-width of ITD-tuning functions in response to transposed tones (0.25 cycles). However, the median half-width of ITD-tuning functions from low-CF neurons in response to Gaussian noise was 0.35 of the period of the CF. The accuracy of the estimated dominant f_m is explored in Section 6.3.4.

Figure 6.6D displays the half-widths of ITD-tuning functions as a function of the CF of the neuron. For low-CF neurons (shown in green), the relationship between CF and half-width was close to that expected if the half-width was a constant proportion of the period of the CF but it deviated from this at both the lowest and highest CFs. The half-widths of neurons with CFs less than 600 Hz were a smaller proportion of the period of CF (median = 0.32) than the half-widths of neurons with CFs greater than 1 kHz (median = 0.52). ITD-tuning functions of neurons with CFs < 600 Hz may be relatively narrow due to tighter phase-locking at low-frequencies (Joris *et al.*, 1994a; Louage *et al.*, 2005). ITD-tuning functions of neurons with CFs above ~1 kHz may be relatively broad if they are strongly influenced by frequencies lower than CF, where phase-locking is tighter and firing rates are more modulated in response to ITDs.

The half-widths of ITD-tuning functions obtained from high-CF neurons in response to broadband transposed noise covered a wide range and showed no significant dependence on CF (linear regression, p = 0.29). This was despite the fact that the f_m presented to each neuron was dependent on CF; the cut-off of the low-pass modulating noise was CF/5. The lack of dependence of half-width on CF in response to broadband transposed noise suggests that the ITD-tuning was determined by a neuron's frequency-filtering properties and was not dictated by a limited range of f_m in the stimulus.

6.3.3. ITD-tuning in response to transposed noise with lower f_m

Nine high-CF neurons (including 6 that contributed to data in Sections 6.3.1 & 6.3.2) were sensitive to ITDs in a transposed noise stimulus with a low-pass cut-off of the modulating noise less than CF/5. Further details and examples of the ITD-tuning functions obtained are given in Section I.2.4. In Figure 6.7, the half-widths of all ITD-tuning functions are shown as a function of the cut-off of the low-pass modulating noise. The half-widths of ITD-tuning functions in response to broadband transposed noise are shown in red and include data obtained at different intensities from the same neuron (see Section I.2.2). There was no significant relationship between the half-width of ITD-tuning functions in response to *broadband* transposed noise and the low-pass cut-off of the neuron (as also shown in Figure 6.6D, N.B. the cut-off frequency of the broadband transposed noise was proportional to CF).

The data in Figure 6.7 are consistent with the notion that in response to transposed noise with cut-off equal to CF/5 (shown in red), high-CF neurons respond with an ITD-tuning function half-width which is characteristic of their preferred envelope tuning. However, if the envelope frequencies present in the stimulus are restricted to frequencies lower than ~400 Hz, the width of the ITD-tuning functions increases, reflecting the lower f_m in the stimulus. 400 Hz may therefore reflect an upper limit above which f_m contribute little to sensitivity to ITDs in high-CF neurons. An upper limit on the f_m at which sensitivity to ITDs can occur is consistent with psychophysical studies using human listeners (Bernstein & Trahiotis, 1994; Bernstein & Trahiotis, 2002).

There was a trend for the narrowest half-width of ITD-tuning functions in response to broadband transposed noise to increase with increasing cut-off frequency (Figure 6.7A; data shown in red). This trend is also present in the data shown in Figure 6.7D. This may reflect a reduction in the f_m at which sensitivity to ITDs can occur as CF increases, consistent with human sensitivity to ITDs (Bernstein & Trahiotis, 2002).

6.3.4. The dependence of half-width on f_m

It was assumed that the half-widths of ITD-tuning functions in response to transposed noise were related to the f_m at which neurons were sensitive to ITDs. This assumption was based on previous data from low-CF neurons, linking the half-widths of ITD-tuning functions in response to broadband stimuli with CF (Yin *et al.*, 1986; McAlpine *et al.* 2001), and from high-CF neurons, linking the half-widths of ITD-tuning functions in response to SAM tones with the f_m of the stimulus (Batra *et al.*, 1993; Joris & Yin, 1995). It was also assumed that a dominant f_m, similar to the CF of low-CF neurons, represented the f_m which dominated sensitivity to ITDs. These assumptions were tested using responses from 8 neurons for which ITD-tuning functions in response to both broadband transposed noise and transposed tones were obtained.

Yin *et al.* (1986) demonstrated that ITD-tuning functions from low-CF neurons in response to broadband noise can be approximated by a composite function. Composite functions are a model of a neuron's response to broadband noise that assumes that sensitivity to ITDs in noise is derived linearly from sensitivity to ITDs in the tones that comprise the noise. It is also possible to create composite curves from response to SAM tones that resemble ITD-tuning functions in response to Gaussian noise (Joris, 1996). In Figure 6.8, composite functions derived from responses to transposed tones (see Section 6.2) are shown in black and ITD-tuning functions generated in response to transposed noise are shown in blue, for 8 neurons for which responses to transposed tones were obtained. The composite functions display remarkably similar shapes to the ITD-tuning functions obtained in response to transposed noise. Troughs and peaks in firing rate occurred at similar ITDs although the maximum firing rates and modulation depths were often different. The composite functions in Figures 6.8C & E are wider than the ITD-tuning functions in response to broadband transposed noise, probably due to over sampling of responses to transposed tones with relatively low-frequency f_m. That composite functions closely

resembled ITD-tuning functions in response to transposed noise suggests, to a first approximation, that ITD-tuning functions in response to transposed noise reflect a neuron's ITD sensitivity at individual f_m.

The estimated dominant f_ms were compared to the f_ms at which sensitivity to ITDs occurred in response to transposed tones, for the same 8 neurons. For 6/8 neurons, the estimated dominant f_m was higher than the maximum f_m at which sensitivity to ITDs occurred in response to transposed tones. Therefore, the estimated dominant f_ms may have over-estimated the f_ms dominating the ITD-tuning functions in response to broadband transposed noise. Note that if dominant f_ms were actually lower than estimated, the data from high-CF neurons in Figure 6.6C (black circles) should be translated down in frequency (to the left). In this case the difference between CF-dependent ITD-tuning of low-CF neurons and f_m-dependent ITD-tuning of high-CF neurons, shown in Figure 6.6C, would be increased.

The synchronisation rate [vector strength x number of spikes (Yin *et al.*, 1986)] was calculated at each f_m as a measure of the sensitivity to ITDs (Figure 6.9). There was a significant linear regression between the estimated dominant f_m and the f_m, in response to transposed tones, at which the synchronization rate was highest ($p = 0.0015$, $R^2 = 0.87$). This may be only weak evidence that the estimated dominant f_m was related to the f_m around which sensitivity to ITDs was greatest; the linear regression was greatly influenced by a trough-type ITD-tuning function with the highest estimated dominant f_m.

6.4. Discussion

The primary objective of the experiments presented in this Chapter was to investigate whether high-CF neurons demonstrate any frequency-dependent tuning to ITDs, similar to CF-dependent tuning to ITDs from low-CF neurons. There was a linear relationship between the peak-ITD of peak-type ITD-tuning functions in response to

a broadband transposed noise and their half-width (Figure 6.5A), reminiscent of the relationship between peak-ITD and half-width observed in the responses of low-CF neurons to broadband Gaussian noise (McAlpine *et al.*, 2001). At low-frequencies, where the half-width of ITD-tuning functions in response to broadband Gaussian noise is proportional to the CF of the neuron, this relationship can be described as CF-dependent ITD-tuning. At high-frequencies, where the half-width of ITD-tuning functions in response to broadband transposed noise is determined by the f_m at which sensitivity to ITDs occurs, this relationship can be described as f_m-dependent ITD-tuning.

The relationship between peak-ITD and f_m-tuning was based on the responses of just 14 neurons with peak-type ITD-tuning functions, 4 of which were modulated with ITD by less than 10 sp/s. It is noted here that neurons sensitive to ITDs in the broadband transposed noise were relatively infrequently encountered. It is likely that the high variability in response rates was partly due to the use of different tokens of noise for each presentation of the transposed noise stimulus and the relatively low number of repeats at each ITD (median = 10). It is also likely that peripheral processing of the sound stimulus, particularly due to its broad bandwidth, resulted in an "internal" stimulus with a small modulation depth and may have contributed to the weakly modulated responses. This is investigated further in Chapter 7.

6.4.1. Comparison with data from neurons with low-frequency CFs

The f_m-dependent ITD-tuning described in response to broadband transposed noise was directly compared to the CF-dependent ITD-tuning of low-CF neurons (Figure 6.6). McAlpine *et al.* (2001) hypothesised that ITD-tuning functions peak at ITDs outside the physiological range so that their steepest slopes occur within the physiological range. It is unlikely that the *peak-ITDs* of peak-type ITD-tuning functions in response to broadband transposed noise could encode for naturally-occurring ITDs because ITD-tuning functions were broad around the peak and peak-

ITDs were often outside the physiological range. It is more likely that the *slopes* of these ITD-tuning functions code for ITDs, although the modulation in firing rate within the physiological range was often small (Figure 6.5B). The small modulation depths may indicate that sensitivity of high-CF neurons to ITDs in a broadband stimulus is not physiologically relevant. In addition, these ITD-tuning functions were obtained in the absence of IIDs that would accompany ITDs in natural stimuli, and are likely to also modulate neural firing rates (Joris, 1996).

The significance of f_m-dependent ITD-tuning is that it suggests that the neural representation of ITDs within both low- and high-frequency sounds is similar. Although previous studies have suggested that sensitivity to ITDs at low- and high-frequencies occurs by similar mechanisms, this is the first study to provide evidence for f_m-dependent ITD-tuning. In the model of ITD-coding proposed by Jeffress (1948), in which ITDs are signalled by the peak firing rates of ITD-sensitive neurons, there is no relationship between the peak-ITD of an ITD-tuning function and the frequency-tuning of the neuron from which it is obtained. The data obtained from high-CF neurons are more consistent with a neural code for ITDs, proposed to account for data from low-CF neurons, which uses the greatest dynamic range of neural firing rates e.g. a model in which two broadly tuned populations of neurons, each from one IC, encode for ITDs at the level of the IC (Figure 1.5.1C).

The peak-ITDs of ITD-tuning functions from both low- and high-CFs neurons occur within a relatively broad range but never fall outside certain limits. At high-CFs, peak (and trough) ITDs occurred within ± 0.5 cycles of the half-width. At low-CFs, peak-ITDs were within 0.5 cycles of the period of CF (McAlpine *et al.*, 2001; Hancock & Delgutte, 2004). This limit on the peak-ITDs that occur is not generally incorporated into models of ITD-sensitivity (Stern & Trahiotis, 1997). Human psychophysical data have been accounted for by assuming the existence of neurons with unphysiologically large peak-ITDs e.g. 2 ms at 500 Hz (van der Heijden & Trahiotis, 1999).

Frequency-dependent ITD-tuning was not identical from low- and high-CF neurons; the peak-ITDs in response to broadband transposed noise occurred at a smaller proportion of the half-width than the peak-ITDs in response to Gaussian noise. The maximum slopes of ITD-tuning functions from low-CF neurons were distributed around 0-μs ITD whereas the maximum slopes were at negative ITDs in response to broadband transposed noise. These differences may be an artefact of the small sample of ITD-tuning functions in response to broadband transposed noise. This explanation is not satisfactory but it is nevertheless noted. It is not considered likely that the difference is due to the use of a transposed noise stimulus rather than a Gaussian noise because peak-ITDs (alternatively, the peak-IPDs) in response to transposed tones and SAM tones were similar.

The differences in f_m- and CF-dependent ITD-tuning (Figure 6.6A-C) may depend on differences in the dominant frequencies (dominant f_m or CF) between high- and low-frequency neurons, and on the half-width as a proportion of the dominant frequency. The dominant frequencies of ITD-tuning functions from neurons with low- and high-frequency CFs are not necessarily equivalent. Low-CF neurons have band-pass tuning to the frequencies at which sensitivity to ITDs occurs. The strong damping of ITD-functions in response to transposed noise, also observed in previously published ITD-tuning functions in response to ITDs in the internal envelope of a Gaussian noise stimulus (Joris, 2003), suggests that tuning to the f_m at which sensitivity to ITDs occurred was not always band-pass. A high-CF neuron may have low-pass tuning for the f_m at which it is ITD-sensitive (see Chapter 3), in which case side-peaks would not be prominent in an ITD-tuning function in response to a broadband stimulus. However, side-peaks may have not been observed because either the ITDs presented were not large enough (6/22 neurons) or because it was not possible to resolve them above the high variation in firing rates (see Figure 6.3). Differences in the dominant frequencies of low- and high- CF neurons do not account for the differences in the ITDs at which the maximum slopes occurred. It is possible that there is an underlying difference, between low-and high-frequencies, in the mechanism that creates a

dependence of peak-ITD on the temporal structure of action potential firing (see 6.3.3).

6.4.2. Comparison with previous physiology

The only published physiological data obtained in response to ITDs in transposed stimuli are presented as part of this thesis in Chapter 3 (Griffin *et al.*, 2005). In response to transposed tones, the cluster of peak-IPDs between -0.1 and 0.3 cycles is consistent with a relationship between peak-ITD and the f_m of the stimulus (see Figure 4.5B).

ITD-tuning functions in response to transposed noise were presented at the Association for Research in Otolaryngology midwinter meeting, 2005 (Endler *et al.*, 2005). These were obtained from the auditory cortex of the cat in response to transposed stimuli with a 1/3 octave noise envelope and a high-frequency tone or narrow-band noise carrier. Endler *et al.* (2005) describe half-widths that were less than 200 μs when the f_m were as low as 64 or 128 Hz i.e. half-widths that were < 0.02 cycles of the period of the f_m. Half-widths in response to low-CF neurons reported here had a median of 0.35 cycles and the half-widths of ITD-tuning functions in response to transposed tones had a median of 0.25 cycles. In addition, Endler *et al.* (2005) report that peak-ITDs were within the physiological range for the cat. This was different from ITD-tuning functions from both low- and high-CF neurons observed in this study, which often peaked outside the range of ITDs that encompass the physiological range experienced by a guinea pig. A possible explanation for the narrow ITD-tuning functions obtained from the cortex of the cat was that the ITD-tuning functions were dominated by onset responses which only occurred around the peak-ITD. In addition, differences between responses in the IC and cortex, species differences or differences in anaesthetic cannot be ruled out as explanation for the different data obtained.

Crow *et al.* (1980) first document that high-frequency neurons show firing rate modulation in response to bands of noise. One previous study (to date) has investigated the sensitivity of high-CF neurons to ITDs in a broadband noise and the peak-ITDs were not reported (Joris, 2003). ITD-tuning functions were recorded from neurons in the IC and were also modelled from the response of ANFs. The envelope of a broadband Gaussian noise, after peripheral processing, depends on the frequency-vs.-intensity response properties of an ANF. The higher the CF, the wider the frequencies an ANF responds to and the faster are the fluctuations of the "internal" stimulus envelope. Therefore, in the modelled ITD-tuning functions there was a relationship between CF and the width of simulated ITD-tuning functions [see also Louage *et al.* (2004)]. However, the half-widths of ITD-tuning functions obtained from neurons with CFs above 2 kHz were not dependent on CF, a similar finding to that described in response to broadband transposed noise. This is consistent with an upper limit in the f_m at which sensitivity to ITDs in the envelope of high-frequency sounds can occur.

The half-widths of ITD-tuning functions in response to the "internal" envelope of a Gaussian noise, reported by Joris (2003), were narrower than the tuning widths in response to transposed noise reported in this study. This may reflect a species difference between cats and guinea pigs in the f_m to which phase-locking occurs. The half-widths of ITD-tuning functions in response to transposed noise in this study were comparable to the half-widths of ITD-tuning functions of neurons with CFs less than 500 Hz in the same species (see Figure 6.6D).

Batra *et al.* (1993) describe an extensive study of high-CF neurons in the IC of the rabbit in response to ITDs in SAM tones. They show that an IID (favouring the contralateral ear) can push the peak-ITD further towards contralateral leading ITDs at low f_m than at high f_m. This could accentuate the relationship between peak-ITD and half-width when an IID and ITD are present. The firing rates of high-CF neurons have been described as being dominated by IIDs (Joris, 1996). However, at

frequencies around 2-4 kHz, the naturally occurring IIDs are relatively small and, in this range, both ITDs and IIDs may be important in shaping spatial-tuning functions.

6.4.3. What is the mechanism for frequency-dependent ITD-tuning across CFs?

The data presented in this Chapter suggests that there is a frequency-dependent ITD-tuning at both low and high-frequency CFs. It is suggested that this likely occurs as a property of the temporal structure of inputs to coincidence detector neurons, whether this is derived from the fine-structure or the envelope of a stimulus. Data from low-CF neurons in the MSO of the gerbil suggests that fast-acting glycinergic inhibition, which is developmentally regulated, creates ITD-tuning functions that peak at an ITD dependent on the frequency of the monaural inputs (Brand *et al.*, 2002; Seidl & Grothe, 2005). Such a mechanism might also result in f_m-dependent ITD-tuning in high-CF neurons.

Studies from the bat MSO suggest that interactions of excitation and inhibition can create tuning to f_m (Grothe *et al.*, 1997; Grothe & Park, 2000). It could be the case that the same neural interactions that determine preference for envelope frequencies (and hence the width of an ITD-tuning function) also result in the preferred tuning to ITDs. The afferents of coincidence detector neurons show low-pass tuning to f_m (Joris *et al.*, 1994a; Joris & Yin, 1998) whereas tuning to ITDs can be band-pass for f_m in the IC (see Chapter 3). This band-pass tuning might originate at the level of the SOC. Whether sensitivity to ITDs at high-frequencies has a physiological role or not, f_m-dependent ITD-tuning might occur as a by-product of pathways in the auditory brainstem. Under this model, the difference in the frequency-dependent ITD-tuning observed at low- and high-frequency CFs could be due to differences in the tuning to f_m compared to tuning to pure tone frequencies, or due to differences in the relationship between excitatory and inhibitory inputs onto coincidence detector neurons at low- and high-frequencies.

Tuning to ITDs has been proposed to occur as a result of delay lines (Jeffress, 1948). Given the observation of f_m-dependent ITD-tuning, the delay line length would have to depend on the f_m tuning of the coincidence detector neuron. In addition, the peak-ITDs in this study could be large (> 500 μs) and are less likely to be due to a difference in the physical path length of projections from the anterior central cochlear nucleus (aVCN) to the MSO. This does not exclude mechanisms for creating ITD-tuning in high-CF neurons such as a synaptic delays or a stereausis mechanism in which delays are created by systematic mis-matches in the CF of inputs to a coincidence detector (Shamma *et al.*, 1989). However, systematic CF mis-matches have not been noted, to date. ITD-tuning of high-CF neurons could result from processing between the SOC and the IC. A candidate would be ITD-sensitive GABAergic inhibition from the DNLL. However, the peak-ITDs of ITD-tuning functions recorded from low-CF neurons in the presence of agonists and antagonists of GABA are unchanged (Ingham & McAlpine, 2005).

It remains unknown how f_m-dependent ITD-processing of high-frequency stimuli occurs but it is tempting to suggest that the mechanism is similar to that which underlies CF-dependent ITD-tuning at low-frequencies.

6.5 Figures and Tables for Chapter 6

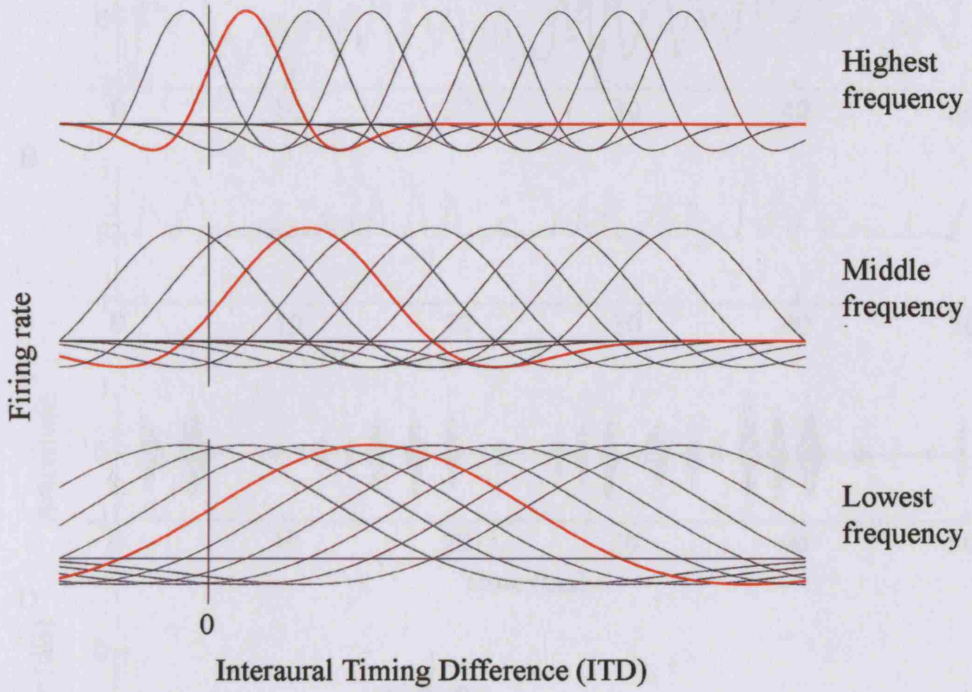

Figure 6.1. Diagrammatic ITD-tuning functions in response to broadband noise from low-frequency neurons with different CFs (highest, middle or lowest). The vertical line marks ITD = zero and peak-ITDs predominantly occur at positive ITDs. *All* ITD-tuning functions represent a Jeffress-type code for ITDs. The *red* ITD-tuning functions represent a code proposed by McAlpine *et al.* 2001, based on the distribution of peak ITDs observed from ITD-tuning functions recorded from the IC.

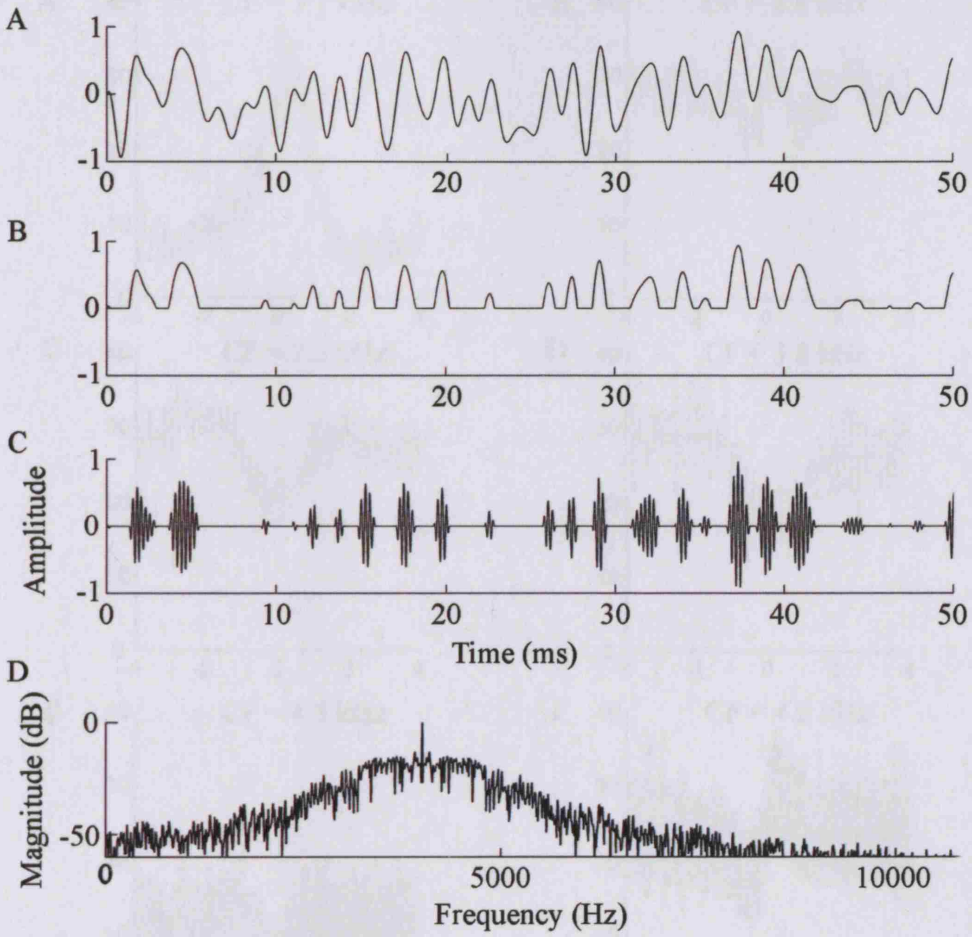

Figure 6.2. An example of a transposed noise waveform. A low-pass (cut-off = 800 Hz), flat spectrum noise (A) was half-wave rectified (B) and multiplied by a high-frequency (4 kHz) carrier tone (C). The spectrum of the resulting stimulus (D) has a broad spread of energy centred on 4 kHz.

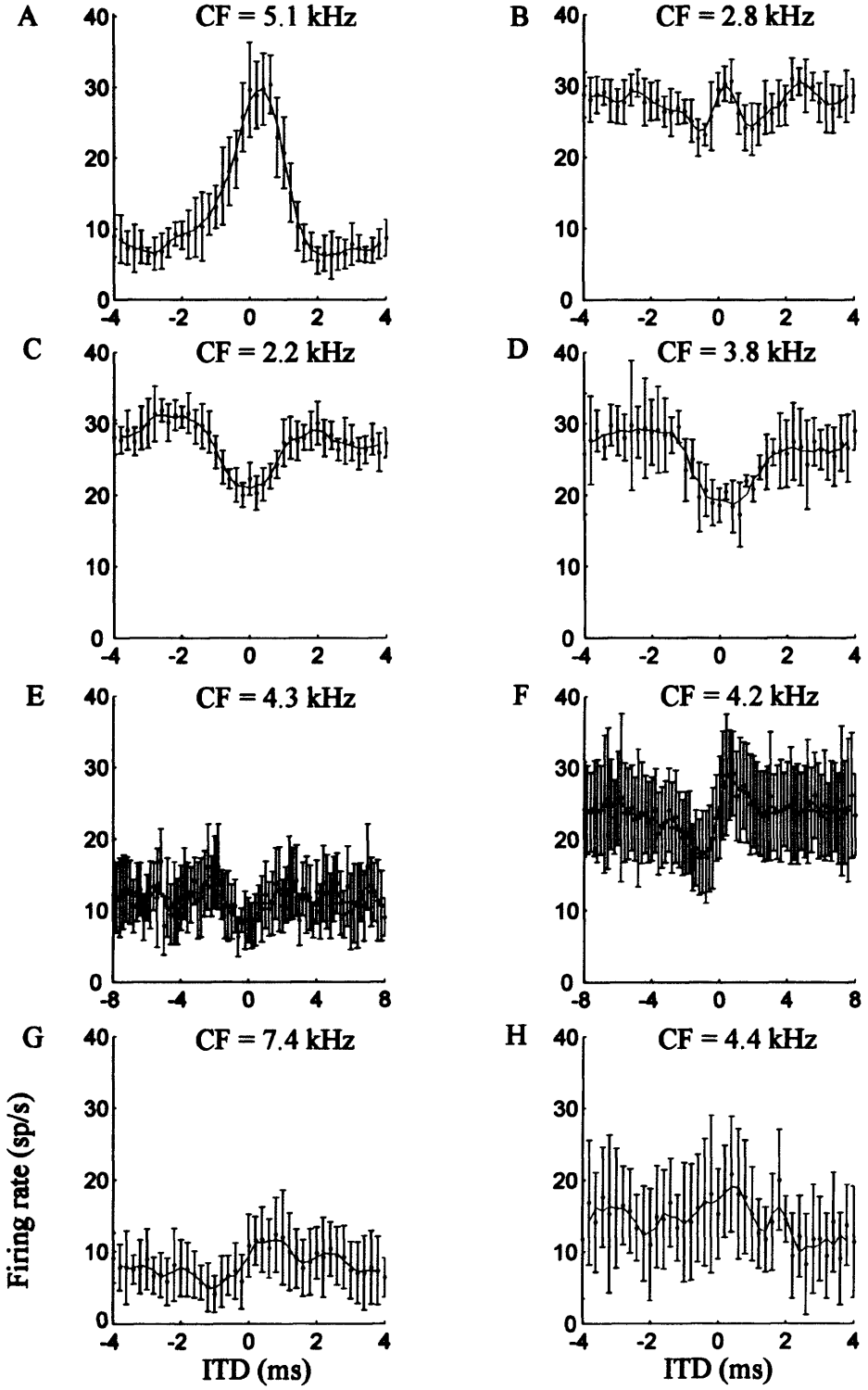

Figure 6.3. See next page for legend.

Figure 6.3. ITD-tuning functions in response to broadband transposed noise. Mean and standard deviation firing rates from 8 neurons: A) 29005 B) 28803 C) 27204 D) 30406 E) 43409 F) 42416 G) 39907 H) 39905. The continuous line marks smoothed firing rate (3-point average). A, B, F, G and H were considered Peak-type. C, D and E were considered trough-type.

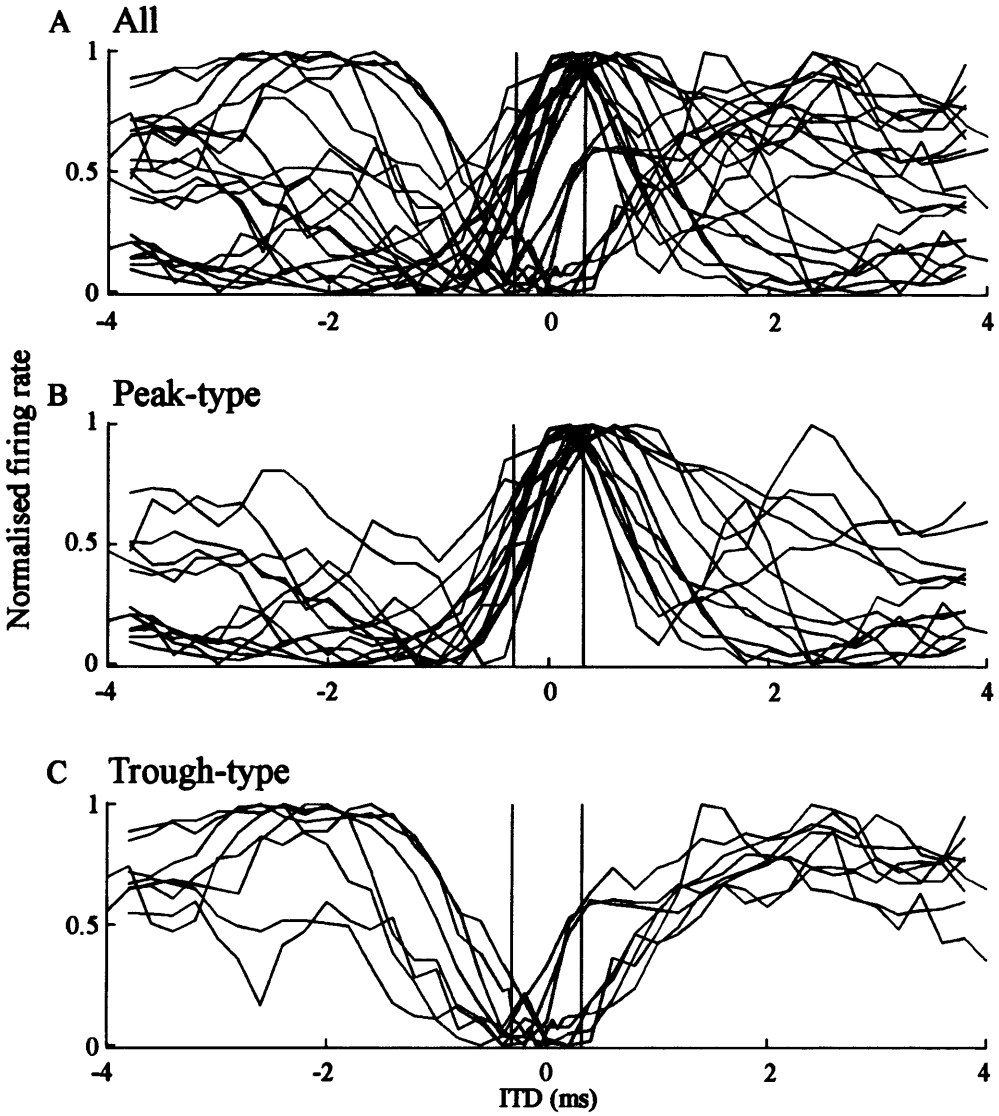

Figure 6.4. Normalised firing rates from 22 neurons, in response to broadband noise, which were modulated in response to ITDs. A) All ITD-tuning functions, B) ITD-tuning functions classified as peak-type (14 neurons) and C) ITD-tuning functions that were classified as trough-type (8 neurons). Firing rates are normalised from 0 to 1 to emphasise the shape of tuning to ITDs. Shaded regions demarcate the possible ITDs that a guinea pig might experience (330 μs; Sterbing *et al.* 2003).

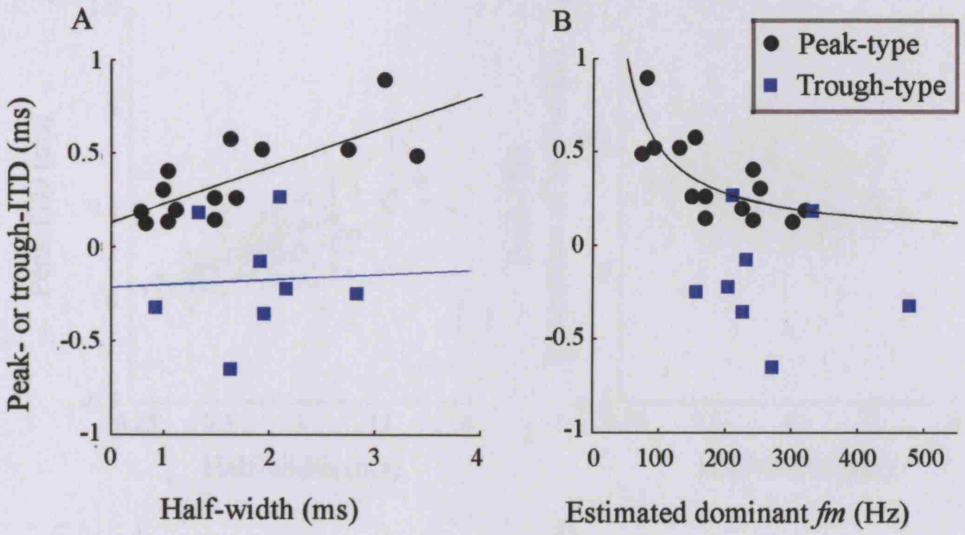

Figure 6.5. The relationship between the peak-ITD (black) or trough-ITD (blue) and the half-width of ITD-tuning functions in response to broadband transposed noise. A) Peak-ITDs and trough-ITDs as a function of half-width. The lines are linear regression fit to the data. Only the regression between peak-ITD and half-width was significant ($p = 0.0018$, $R^2 = 0.57$). (B) Peak-ITD and trough-ITD as a function of the estimated dominant f_m. The line is the same significant linear regression shown in A.

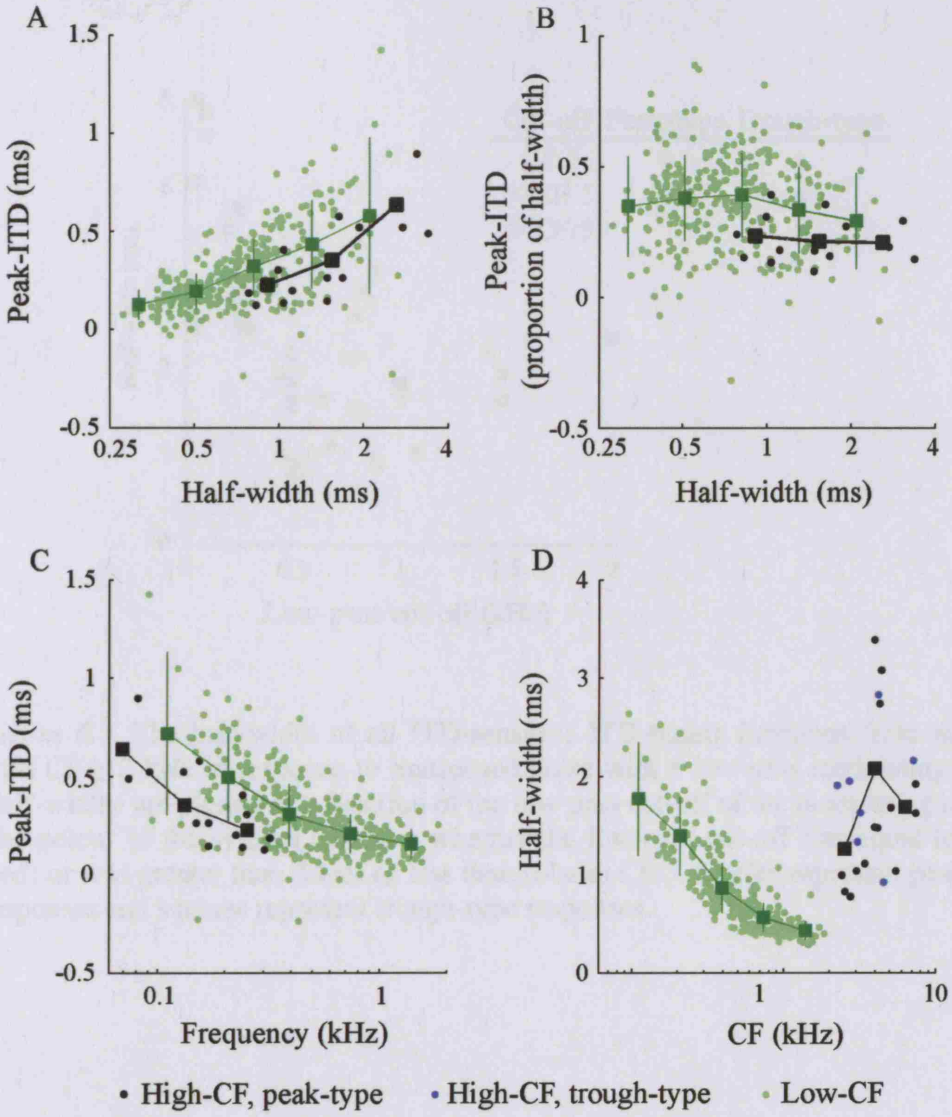

Figure 6.6. ITD-tuning functions in response to a Gaussian noise recorded from low-frequency neurons (green) were compared to ITD-tuning functions in response to broadband transposed noise which were classified as peak-type (black) or trough-type (blue). A) Peak ITD as a function of half-width. B) Peak ITD (as a proportion of the half-width). C) Peak ITD as a function of CF or estimated dominant f_m. D) Half-width as a function of CF. Circles mark data from individual neurons. Squares mark mean data after grouping into 3 or 5 bins. For low-CF neurons bins were logarithmically spaced between half-width = 0.25 and 2.65 and CF = 0.08 - 1.9. For high-CF neurons bins were logarithmically spaced between half-width = 0.7 – 3.4 ms, dominant f_m = 0.05 and 0.035 and CF = 2 and 8 kHz. Error bars (± standard deviation) are only shown for the data obtained from high-CF neurons.

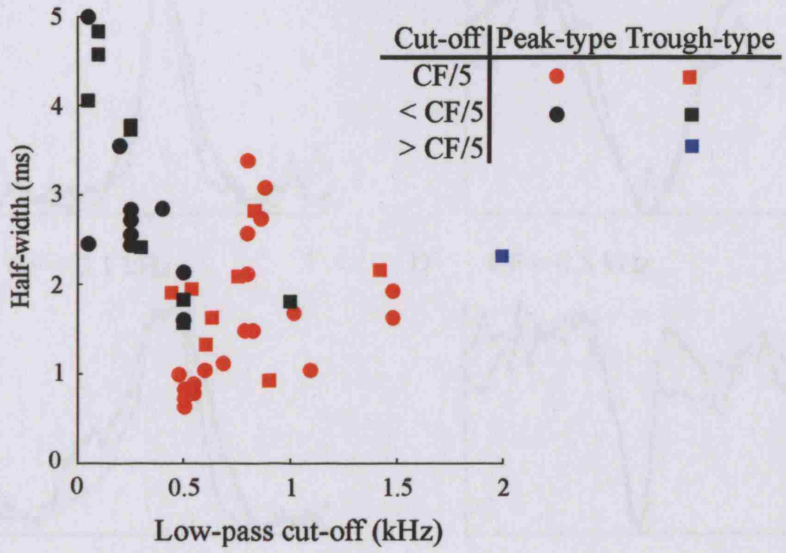

Figure 6.7. The half-width of all ITD-sensitive ITD-tuning functions from neurons with CF $\geq$ 2 kHz in response to transposed noise with a low-pass modulating noise. Half-widths are plotted as a function of the low-pass cut-off of the modulating noise. The colour of the symbol indicates whether the low-pass cut-off was equal to CF/5 (red) or was greater than (blue) or less than (black) CF/5. Circles represent peak-type responses and squares represent trough-type responses.

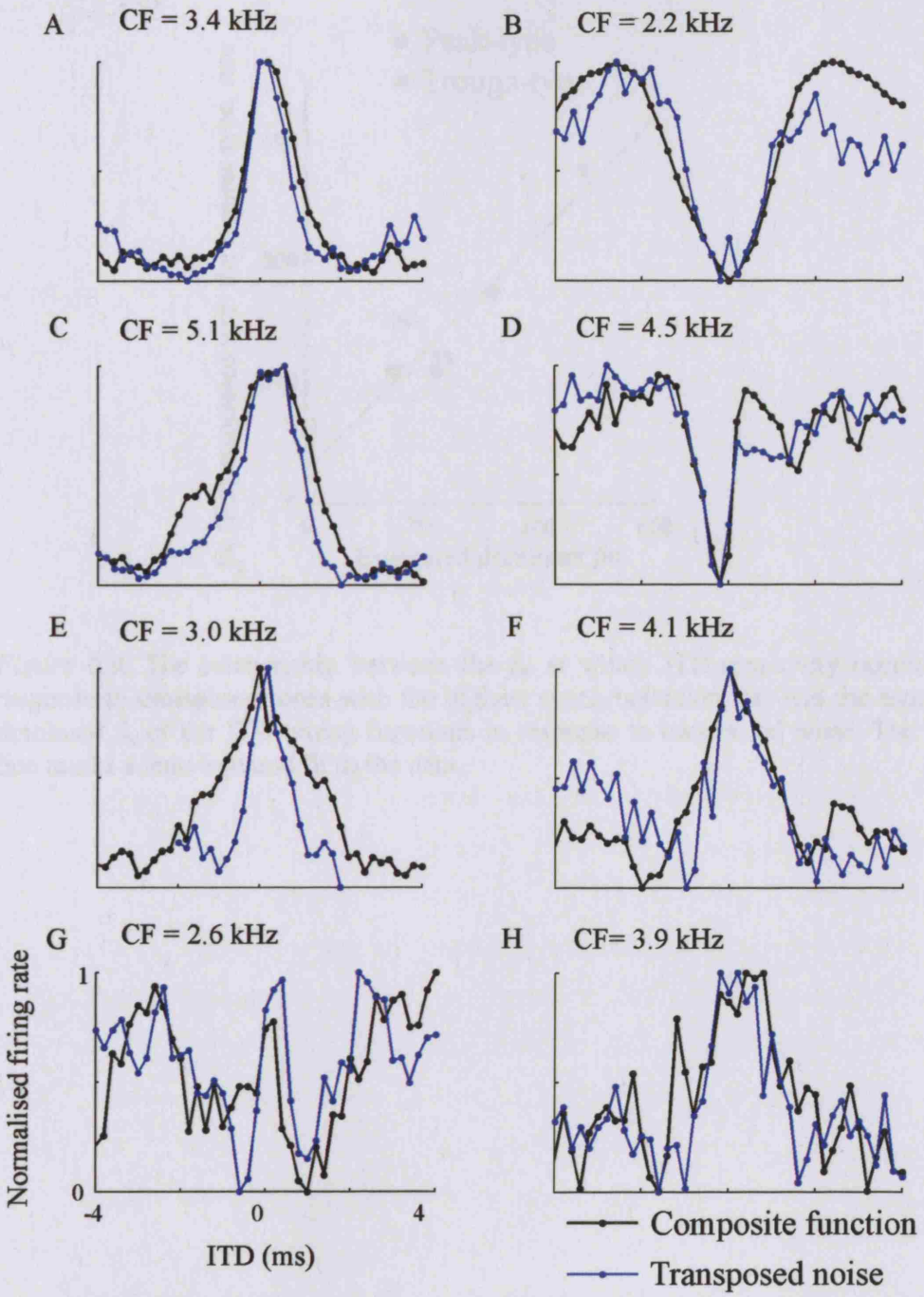

Figure 6.8. ITD-tuning functions (normalized between 0 and 1) in response to broadband transposed noise (blue) and composite functions derived from firing rates in response to transposed tones (black). Neurons and f_m from which composite curves were derived are noted in Section 6.2.

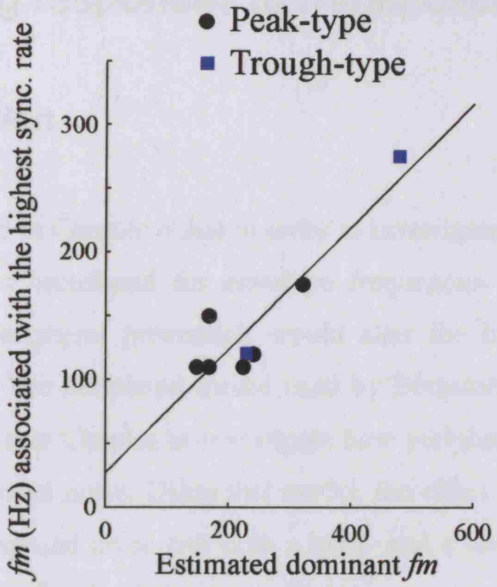

Figure 6.9. The relationship between the f_m at which ITD-sensitivity occurred in response to transposed tones with the highest synchronization rate and the estimated dominant f_m of the ITD-tuning functions in response to transposed noise. The black line marks a least-squared fit to the data.

7. Modelling responses to transposed noise stimuli

7.1 Introduction

It was emphasised in Chapter 6 that in order to investigate f_m-dependent ITD-tuning a stimulus that was broadband for envelope frequencies was required. It was also suggested that peripheral processing would alter the broadband transposed noise stimulus chosen. The peripheral model used by Bernstein & Trahiotis (2002, 2003) was employed in this Chapter to investigate how peripheral processing might alter a broadband transposed noise. Using this model, the effect of peripheral processing on a broadband transposed noise and both a high- and a low-frequency band-pass flat-spectrum noise (see Section II.1) are compared.

7.2. Results

A 50-ms portion of three noise waveforms which were compared are shown in Figure 7.1A-C. The transposed noise waveform was centred on 4 kHz with the low-pass cut-off of the low-pass modulating noise equal to 800 Hz (see also Figure 6.2.1). This is equal to the broadband transposed noise stimulus presented to a neuron with a CF of 4 kHz in the experiments described in Chapter 6. A high-frequency flat-spectrum noise waveform, with a band-width of 1.6 kHz, centred on 4 kHz, is shown in Figure 7.1B. This was chosen because, as in the transposed noise waveform, its spectral components were centred on 4 kHz. Also equivalent to the transposed noise waveform, its spectral components were broader than the width of the effective "filter" or spectral receptive field of an ANF with CF = 4 kHz (Evans *et al.*, 1992). A low-frequency flat-spectrum noise waveform, with a band-width of 800 Hz, centred on 500 Hz, is shown in Figure 7.1C. 500 Hz was chosen as a low-frequency region at which to compare the processing of low- and high-frequency noise stimuli. The band-width of the noise was larger than the width of the effective "filter" or spectral receptive field of an ANF with CF = 500 Hz (Evans *et al.*, 1992). In all three

examples, therefore, it was predicted that peripheral processing would alter the noise waveforms.

The two high-frequency noise waveforms were filtered with a gamma-tone filter centred at 4 kHz, to model the spectral receptive field of an ANF with CF = 4 kHz. The results are shown in Figures 7.1D & E. The resulting waveforms were smoothed compared to the original waveform because high-frequency fluctuations in their envelopes were removed by the filtering of frequencies distant from 4 kHz. The off-periods in the transposed noise stimulus were virtually absent after filtering by a gamma-tone filter. Figure 7.2.1F shows the low-frequency noise waveform filtered through a gamma-tone filter centred on 500 Hz. The 500 Hz-filter was relatively narrow, compared to the high-frequency filter, and the resulting low-frequency waveform was dominated by frequencies around 500 Hz.

Panels G, H & I of Figure 7.1 show the noise waveforms after filtering by the peripheral model of Bernstein & Trahiotis (2002). After filtering with a gamma-tone filter, the waveforms were square-law rectified, envelope compressed and low-pass filtered. The resulting waveforms are a model of the effective internal stimulus. They can also be thought of as a model of action potential firing in an ANF. There were no distinct off-periods in the waveform derived from a transposed noise (Figure 7.1G) but the waveform did have a greater modulation depth than the waveform derived from a high-frequency flat-spectrum noise (Figure 7.1H). Figure 7.1I shows the result of the peripheral model applied to the low-frequency noise. In this case, the peripheral model was altered by omission of the 150-Hz low-pass filter that modelled the reduction in sensitivity to f_m observed at high-frequencies. The 500-Hz fluctuations in the waveform were preserved and the modulation depth of the stimulus remained high. Broadband transposed noise, unlike transposed tones, does not mimic the representation of low-frequency noise stimuli within an ANF. This is because the filtering removes some of the spectral information in the transposed noise, effectively removing the "off-periods". However, compared to a high-frequency flat-spectrum noise, the broadband transposed noise does have a greater

depth of modulation and it is therefore predicted that the modulation depth of ITD-tuning functions would be greater in response to transposed noise as compared to the more conventional noise stimulus.

To investigate whether it was likely that a broadband transposed noise would provide an advantage in detecting ITDs over the high-frequency flat-spectrum noise, the normalised correlation was calculated from 10 independent tokens of a transposed noise and the high-frequency flat-spectrum noise at each of 49 ITDs between 20 μs and 1000 μs. The normalised correlation was also calculated for a SAM and transposed tone centred at 4 kHz and with f_m = 125 Hz. All waveforms were subject to the same peripheral model (which included gamma-tone filtering, low-pass filtering, square-law rectification and compression) before calculation of the normalised correlation. The results are displayed in Figure 7.2A, where the mean normalised correlation at each ITD is plotted for the noise stimuli. A dotted horizontal line marks a constant level of interaural correlation, demonstrating that different ITDs are required to achieve the same level of decorrelation (of the effective internal stimulus) for each stimulus. Broadband transposed noise generated a faster reduction in the normalised correlation than flat-spectrum noise, as ITD was increased from 0 μs. However, the normalised correlation derived from both SAM and transposed tones decreased even more quickly as ITD increased from 0 μs. For a neuron with CF equal to the f_c of the stimuli, these results suggest that neural discrimination of ITDs from 0-μs ITD would require a larger change in ITD in response to a broadband transposed noise than an SAM or transposed tone with f_m = 125 Hz. However, it suggests that neural discrimination of ITDs would be more sensitive in response to a broadband transposed noise than a more conventional high-frequency noise, for neurons with CF equal to the f_c of the stimuli. Note that these results might suggest that ITDs in a broadband transposed noise stimulus are more difficult to detect than in a transposed tone with 125 Hz f_m. However, this was not the case (personal observation, data not shown). Responses of neurons with CFs off the f_c of the stimulus may be involved in the perception of the stimulus.

In the experiments described in Chapter 6, independent tokens of transposed noise were presented. Therefore, there was variation in the correlation of the waveforms and in the *rms* intensity of the waveforms on different presentations. To investigate the variability in an ITD-tuning function in response to transposed noise that could be attributed to the stimulus, the cross-product was calculated, after peripheral processing, of 10 independent tokens of transposed noise, each at a range of ITDs (± 16 ms, in 200 μs steps). The peripheral model was the same as that used in Bernstein & Trahiotis (2002) and the transposed noise was centred on 4 kHz, with an 800-Hz low-pass modulating noise. The resulting mean cross-products (± standard deviations) are shown in Figure 7.2B, as a function of ITD. The modelled ITD-tuning function was broad with a half-width around 4 ms, at the upper range of the half-widths of the neural ITD-tuning functions (Figure 6.5). The modelled ITD-tuning function was also somewhat noisy with the standard deviation at each ITD of a similar magnitude to the modulation depth of tuning to ITDs. The dc offset of the model ITD-tuning function was also high compared to its modulation depth. Noise tokens for this model were 635 ms long (see Section 2.2.). The variation would be less with a longer stimulus or greater with a shorter stimulus. In the physiological experiments presented in Chapter 5, stimuli were either 200 or 1000 ms in duration. Variation in firing rates would also be expected due to internal, "neural" noise in the auditory system and possibly due to changes in anaesthetic state and the history of stimulus presentation.

7.3. Discussion

A mathematical model was used to investigate the likely effect of peripheral processing on the broadband transposed noise stimulus used in the experiments described in Chapter 6. The peripheral model was the same as that used in Bernstein & Trahiotis (2002) and was also used to investigate the influence of peripheral processing on the shapes of ITD-tuning functions in response to transposed and SAM

tones (see Section 4.3.2). The results suggest that the effective internal modulation depth of the transposed noise stimulus was small compared to that of a high-frequency transposed tone or to a low-frequency flat-spectrum noise. This, and the noise attributable to the use of independent tokens of a transposed noise (as well as "internal" noise inherent to the neurons), can explain the small modulation depths and high variation in firing rates of ITD-tuning functions obtained in response to broadband transposed noise (see Chapter 6).

A transposed noise with a bandwidth such that most of its spectrum falls within an auditory filter can, in principle, provide a similar pattern of action potential firing in a high-CF ANF to that seen in a low-CF ANF in response to a narrow-band noise (Bernstein & Trahiotis, 2003). In order to mimic action potential firing in a low-CF ANF, a transposed noise should be band-pass for f_m. This is because the responses of low-CF ANFs can be considered to act as a band-pass filter for pure tone frequencies. High-frequency IC neurons do not necessarily have band-pass tuning for envelope frequency in terms of their rMTFs (see Section I.1.4) or in terms of the f_m at which they are sensitive to ITDs (see Chapter 3). Therefore a stimulus broadband for f_m was required to examine the preferred tuning to f_m of IC neurons in the experiments presented in Chapter 6. The broadband transposed noise fulfilled this criterion. However, ITD-tuning functions obtained were weakly modulated. The weak modulation was due to peripheral filtering which removed frequencies outside the spectral receptive field of ANFs with CF equal to the CF of the stimulus. This removes the "off-period" characteristic of transposed stimuli which are dependent on the five central frequency components that generate each frequency in the envelope of the stimulus. In the case of a broadband transposed noise centred at 4 kHz, 800 Hz was the highest f_m in the envelope. The five central frequency components of 800 Hz modulation are 4 kHz are at 2.4, 3.2, 4.0, 4.8 and 5.6 kHz. These components, other than the one at 4 kHz, will be heavily attenuated by an ANF with CF = 4 kHz (Evans *et al.*, 1992). The amplitude of the component of the spectrum of the broadband transposed noise at 4 kHz, after peripheral filtering, was therefore a larger proportion of the spectrum after peripheral filtering than beforehand. This generates the dc offset

in the waveform, also observed in the model ITD-tuning function shown in Figure 7.2B.

Chan *et al.* 1987 reported the study of a relatively high-CF neuron [CF = 1300 Hz; see Figure 16 of Chan *et al.* (1987)] which was sensitive to ITDs at pure tone frequencies below ~1300 Hz. The neuron responded to pure tones of higher frequency but was not ITD-sensitive at these frequencies. These higher frequencies, in a composite curve, caused a dc shift in the ITD-tuning function compared to the ITD-tuning function in response to a low-pass noise that excluded frequencies to which the neuron was not sensitive. The presence of f_ms, in a transposed noise stimulus, to which a neuron was not sensitive to ITDs would likely have a similar effect in that ITD-tuning functions would show an increased dc component and a reduced modulation depth. This would be increasingly likely to occur at the highest CFs investigated in the experiments described in Chapter 6 because spectral receptive fields widen with increasing CF but the maximum f_m at which sensitivity to ITDs occurs may decrease. This may have biased against finding neurons with relatively high CFs that were sensitive to ITDs in broadband transposed noise.

A stimulus with $f_m \leq 400$ Hz is broadband in terms of the f_m to which a high-CF neuron is likely to be sensitive to ITDs (see Chapters 3 and 6). A transposed noise stimulus constructed from a low-pass noise with a cut-off of 400 Hz, independent of CF, would likely have been broad enough to investigate frequency-dependent ITD-tuning and to provide an effective internal stimulus with a greater modulation depth than the broadband transposed noise used in the experiments described in Chapter 6. However, neurons with the lowest CF would then filter a greater proportion of the stimulus than neurons with higher CFs and wider filter widths. This may bias against finding neurons with relatively low CFs (closest to 2 kHz) that are sensitive to ITDs in such a stimulus.

The broadband transposed noise stimulus used in Chapter 6 was appropriate for investigating f_m dependent ITD-tuning because the ITD-tuning functions obtained

reflected neurons' preferred tuning to f_m and ITDs. However, there was a trade-off between finding a stimulus which was sufficiently broadband in f_m to investigate neurons' preferred tuning to envelope ITDs and using a stimulus that generated strong sensitivity to ITDs within the spectral receptive field of a single auditory neuron. The results presented in this Chapter suggest that the broadband transposed noise stimulus used in the experiments in Chapter 6 provided little advantage over a more conventional broadband noise stimulus. A transposed noise stimulus with a fixed low-pass cut-off of 400 Hz may have been sufficiently broadband to examine neurons' preferred tuning to ITDs.

The modelled ITD-tuning function (Figure 7.3B) had a wide half-width which was shaped by the low-pass filter for f_m included in the peripheral model. The half-width of ITD-tuning functions is likely to be determined by a neuron's tuning to ITDs at different f_m (see Chapter 6). The narrower ITD-tuning functions observed in neuronal responses could be due to different tuning to f_ms than captured by the low-pass filter in the model. For example, neural tuning to f_m can be band-pass (e.g. see Figure 3.5). Narrower ITD-tuning functions could also be the result of some processing of ITD-tuning functions that occurs after the level of the SOC (Fitzpatrick *et al.*, 1997; Fitzpatrick *et al.*, 2002; Ingham & McAlpine, 2005).

7.4 Figures and Tables for Chapter 7

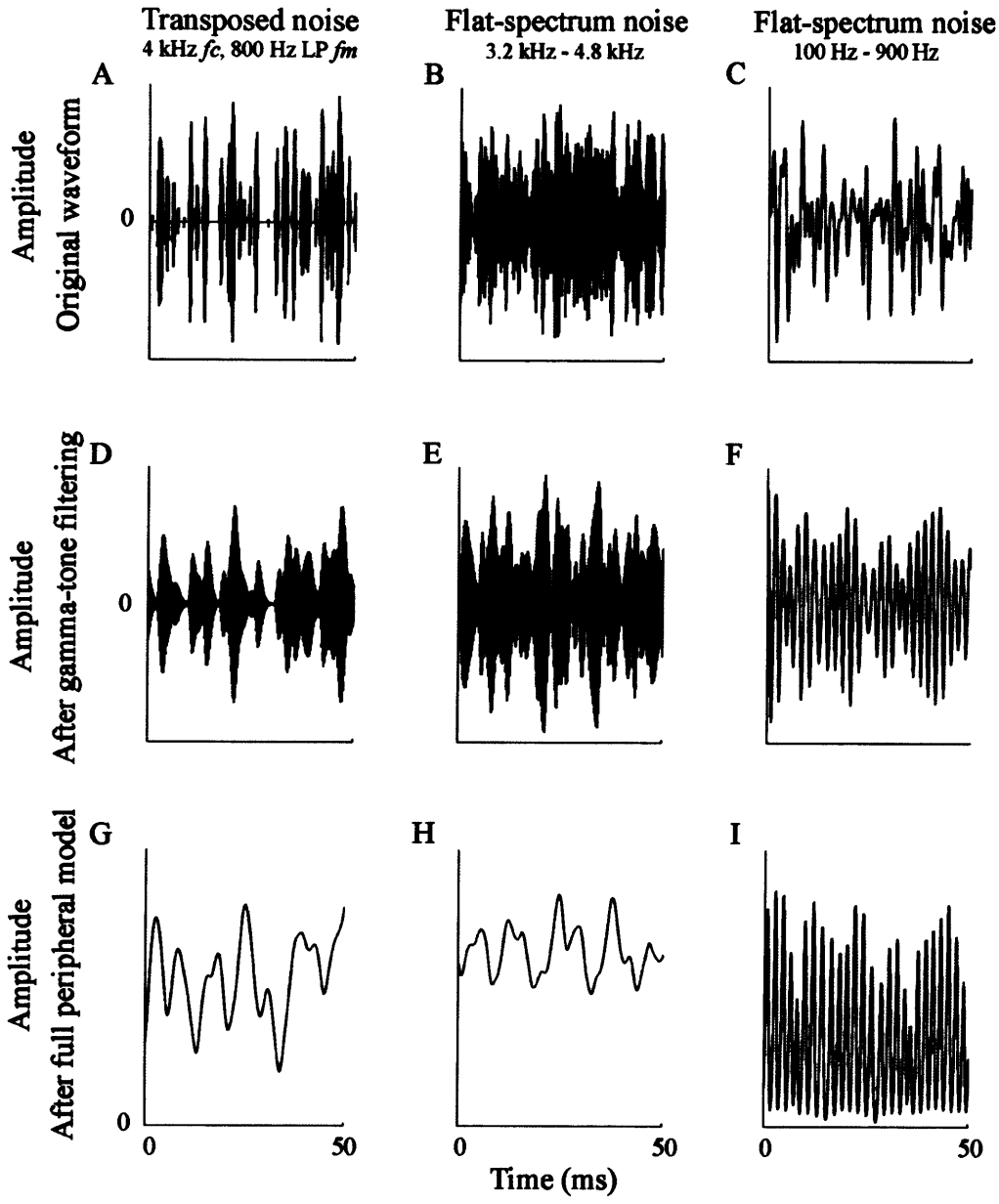

Figure 7.1. The effect of the model of peripheral filtering on three different noises. A) Tranposed noise with low-pass cut-off equal to $f_c/5$, B) 1.6 kHz wide flat-spectrum noise centred on 4 kHz and C) 800 Hz wide flat-spectrum noise centred on 500 Hz.

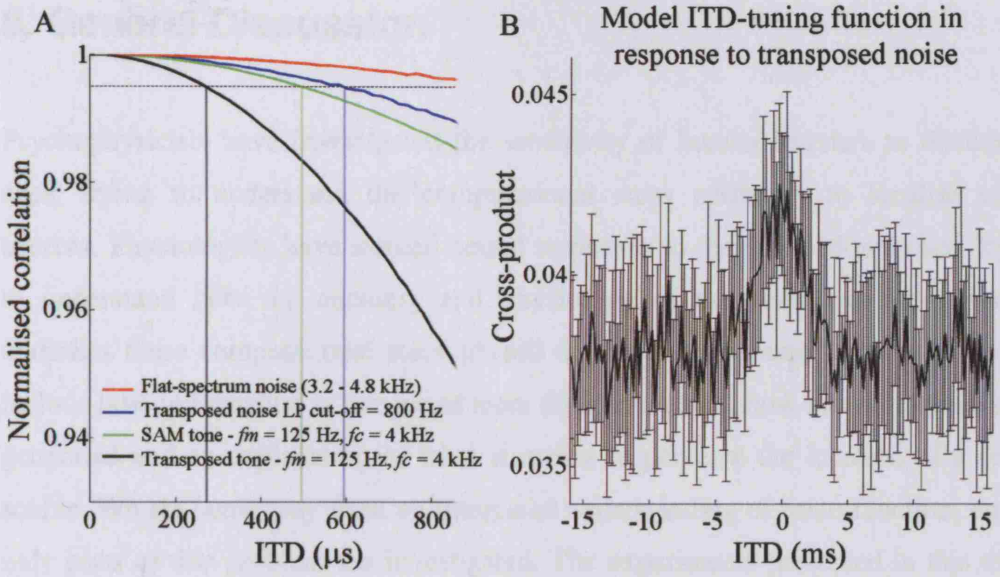

Figure 7.2. Model ITD-tuning functions. A) Thick black, green, blue and red lines show the reduction in the mean normalised correlation that occurred as ITDs increase from 0 µs of modelled responses to transposed tones, SAM tones, transposed noise centred at 4 kHz and with the low-pass cut-off of the modulating noise equal to 800 Hz and a flat-spectrum noise band-pass between 3.2 and 4.8 kHz. The horizontal dotted line marks a constant level of the normalised correlation and vertical lines mark the ITD at which the 4 curves reach that value of normalised correlation. B) A model ITD-tuning function in response to broadband transposed noise.

8. General Discussion

Psychophysicists have investigated the sensitivity of human listeners to directional cues, trying to understand the computational steps necessary to localise sound sources. Physiologists have studied neural responses to these directional cues, trying to understand how the anatomy and physiology of the central auditory system underlies these computational steps. A full description of sound localisation would include how information is integrated from different senses, how motor responses are generated and an explanation of what it means to perceive the location of a sound source. We are some way from an integrated understanding of brain function, and so only parts of this problem are investigated. The experiments presented in this thesis provide evidence for how ITDs in the envelope of high-frequency sounds are encoded.

Neural sensitivity to ITDs was greater in response to transposed tones as compared to SAM tones: more neurons were sensitive to ITDs, JNDs for detection of an ITD were lower and the modulation depths of ITD-tuning functions were greater in transposed tones as compared to SAM tones. Neural sensitivity to ITDs in transposed tones, measured by JNDs, could be equivalent to sensitivity of low-CF neurons to ITDs in low-frequency pure tones, consistent with psychophysical findings. Consideration of sensitivity to ITDs within the envelope of high-frequency sounds led to an interest in the salience of onset-ITDs. Psychophysical experiments were designed to investigate the salience of an onset-ITD within a high-frequency stimulus. The results suggest that onset-ITDs within a high-frequency sound are useful when ongoing-ITDs are ambiguous, otherwise ongoing-ITDs are dominant. The peak-ITDs of ITD-tuning functions in response to broadband transposed noise were dependent on the f_m at which neurons are sensitive to ITDs. This was reminiscent of the CF-dependent ITD-tuning of low-CF neurons.

In this Chapter the results presented in Chapters 3-7 will be discussed, along with the findings of previous studies, under four headings: 1) the similarities and differences between sensitivity to ITDs at high- and low-frequencies, 2) the mechanism underlying improved sensitivity to transposed tones compared to SAM tones, 3) the duplex theory and the physiological relevance of ITDs with the envelope of high-frequency sounds and 4) the potential clinical relevance of the studies presented in this thesis.

8.1 The similarities and differences between sensitivity to ITDs at low- and high-frequencies

ITDs can be discriminated when the ITD of the carrier or envelope of a low-frequency, complex sound is changed (Bernstein & Trahiotis, 1985a). Discrimination of ITDs in high-frequency complex sounds only occurs if the ITD in the envelope of the sound is changed; listeners are insensitive to changes in carrier ITD within high-frequency complex sounds (e.g. Henning & Ashton, 1981). This difference is due to an upper limit in the frequency at which stimulus-locking of action potentials occurs in ANFs, and also due to the ability of second order afferents and coincidence detector neurons to follow the temporal structure of action potential firing in ANFs (Joris *et al.*, 1994a; Joris, 2003). Phase-locking in response to the carrier of a complex waveform, or a pure tone, can only occur in low-CF ANFs (Johnson, 1980; Palmer & Russell, 1986). Stimulus-locking in high-CF ANFs follows the envelope of a complex sound, but not the carrier (Joris & Yin, 1992). The envelope structure can be present within the sound stimulus, or can be created by peripheral processing and therefore be present within the "internal" stimulus (Joris, 2003).

ITD-tuning functions in response to both low- and high-frequency sounds have been described as being dominated by a peak in firing rate, a trough in firing rate or described as containing both a prominent peak and a prominent trough (Yin *et al.*, 1986; Batra *et al.*, 1997a; McAlpine *et al.*, 1998; McAlpine *et al.*, 2001). The

proportion of ITD-tuning functions that are classified as trough-type as compared to peak-type is higher amongst neurons with high-frequency CFs (Batra *et al.*, 1993; Griffin *et al.*, 2005). This is likely to reflect the greater number of neurons in the LSO, compared to the MSO, with high-frequency CFs (Guinan *et al.*, 1972). Joris suggests that there are no neurons > 6.1 kHz in the MSO that are sensitive to ITDs in the envelope of high-frequency sounds (Joris, 2003), although ITD-sensitivity has been recorded from a neuron with CF equal to 9 kHz in the cat MSO (Oliver *et al.*, 2003). There are also neurons in the LSO with CF > 6 kHz that generate sensitivity to ITDs (Joris & Yin, 1995). However, it is likely that the number of high-CF neurons sensitive to ITDs decreases with increasing CF, as does the f_m at which sensitivity can occur (Chapters 3 & 6; Joris, 2003).

In the IC, ITD-tuning functions from both low- and high-CF neurons have peak-ITDs that predominantly occur at contralateral leading ITDs (Yin *et al.*, 1986; Kuwada *et al.*, 1987; Batra *et al.*, 1993; McAlpine *et al.*, 2001). This is likely to reflect the ipsilateral projection of MSO neurons (Oliver, 2000) which usually respond maximally to sounds leading in time at the contralateral ear (Yin & Chan, 1990). LSO neurons respond maximally to sounds originating from ipsilateral space. LSO neurons send excitatory projections to the contralateral IC and are likely to contribute to the preference for contralateral sounds in the IC (Park, 1998). Lesion studies support the notion that the left inferior colliculus represents right space and *vice versa* (Jenkins & Masterton, 1982; Kelly & Kavanagh, 1994). This may be problematic for models of ITD-coding in which ITD-sensitivity occurs from comparison of broadly tuned neurons which form a "channel" in each IC (McAlpine *et al.*, 2001).

Both peak- and trough-ITDs in response to transposed noise were never more than ± 0.5 cycles of the half-width. Neurons with low-frequency CFs do not have peak or trough-ITDs that are larger than half of the period of CF. Peak-ITDs from low-CF neurons are distributed at 45° IPD (McAlpine *et al.*, 2001) and trough-ITDs from low-CF neurons occur close to 0 μs, with the side-peak which is closest to 0-μs ITD

not exceeding 180° IPD (T. Marquardt – personal communication). The limit at which peak-ITDs occur can also be described as CF- or f_m-dependent ITD-tuning. For neurons with low- and high-frequency CFs, the widest ITD-tuning functions tend to peak at the largest ITDs. This is consistent with a model of ITD-coding, at both low- and high-frequencies, based on broadly-tuned populations of neurons (McAlpine *et al.*, 2001) and not on the peaks of ITD-tuning functions (Jeffress, 1948).

The sensitivity of both low- and high-CF neurons to noise stimuli can be modelled by their sensitivity to pure tones or to single f_ms of a high-frequency carrier (Yin *et al.*, 1986). Composite functions do not predict exactly the shape of an ITD-tuning function in response to broadband stimuli, e.g. the degree of damping is often different and the overall modulation depth or mean firing rate may not be predicted, but their shapes around the central peak or trough are remarkably similar (Chapter 6; Joris *et al.*, 2005). ITD-tuning functions in response to transposed noise were generally highly damped, whereas ITD-tuning functions from low-CF neurons in response to Gaussian noise often show side-peaks and troughs (Joris *et al.*, 2005). Low-CF neurons have band-pass tuning to the frequencies at which they are ITD-sensitive due to the frequency-filtering properties of low-frequency ANFs. High-CF neurons do not necessarily have band-pass tuning to the f_m at which they are ITD-sensitive. Therefore, the tuning to the f_m at which they are ITD-sensitive is not exclusively determined by the frequency-filtering properties of high-frequency ANFs; the shapes of ITD-tuning functions from high-CF neurons are not dependent on CF (Chapter 6).

Improvements in performance in an ITD-discrimination task, when training is carried out using low-frequency sounds, translates to an improvement in performance using high-frequencies, and *vice versa* (Rowan & Lutman, 2005). This suggests that there are processes common to discrimination of both low- and high-frequency ITDs, which can be learnt during training. The role of onset-ITDs also seems to be similar across low- and high-frequencies (Chapter 5; Buell *et al.*, 1991). Onset-ITDs are most

useful when ongoing-ITD cues are ambiguous. Otherwise ongoing-ITDs are most salient.

Sensitivity to ITDs in the envelope of high-frequency sounds has typically been shown to be poorer than in response to low-frequency sounds. However, psychophysical and neural thresholds for detection of ITDs within transposed tones are comparable to thresholds for detection of ITD in low-frequency tones (Bernstein & Trahiotis, 2002; Griffin *et al.*, 2005). This suggests that ITD-sensitivity at high-frequencies is partly limited by peripheral processing, as proposed by Colburn & Esquissaud (1976). Transposed tones overcome this by providing high-CF ANFs with a similar probability of action potential firing as observed in low-CF ANFs in response to low-frequency tones.

Peripheral processing is likely to cause high-frequency ANFs to be weakly modulated in firing rate in response to Gaussian noise, leading to ITD-tuning functions with small modulation depths (Chapter 7; Middlebrooks & Green, 1990). Sensitivity to ITDs within the envelope of high-frequency sounds reduces with decreasing modulation depth of the stimulus (Henning, 1974a; McFadden & Pasanen, 1976; Nuetzel & Hafter, 1981; Stellmack *et al.*, 2005). In contrast, low-frequency ANFs are well-modulated due to the hyperpolarisation of IHCs, which can completely shut off action potential firing in ANFs (Palmer & Russell, 1986).

Human sensitivity to ITDs is more dependent on the intensity of high-frequency sounds than low-frequency sounds (McFadden & Pasanen, 1976; Nuetzel & Hafter, 1976). Dreyer *et al.* (2005) provide evidence the dependence of phase-locking on intensity is different between low- and high-frequency ANFs. The increased dependence on intensity of phase-locking in high-CF ANFs is likely to be explained by differences in the mechanism underlying phase-locking in low- and high-CF ANFs. The stimulus-locking required for sensitivity to ITDs in high-frequency sounds is dependent on the amplitude modulation of the sound. In contrast, stimulus-locking in low-CF neurons is generated by periodic hyperpolarisation and

depolarisation of IHCs, generated due to sensitivity to the fine-structure of the sound. Entrainment to pure tones in the low-CF afferents to the MSO (and LSO) is high and its dependence on intensity has a dynamic range that rises rapidly from threshold to saturation (Joris *et al.*, 1994a).

Neural JNDs for ITD discrimination were similar in response to pure tones around 125 Hz and transposed tones with f_m around 125Hz. However, JNDs in response to pure tones decrease as CF increases above 125 Hz, whereas JNDs in response to transposed tones increase as f_m increases above ~125 Hz (up to ~500 Hz). This can be explained by a central auditory system limit on the f_m at which sensitivity to ITDs occurred; there was no trend for the f_m at which sensitivity to ITDs can occur to increase with increasing CF as predicted from peripheral processing (Chapter 3). This limitation may occur in the LSO (or MSO; Joris & Yin, 1998) and is consistent with an upper f_m limit observed in psychophysical experiments (Bernstein & Trahiotis, 1994).

Psychophysical studies show that the maximum f_m at which sensitivity to ITDs can occur is inversely related to the f_c of the stimulus (Bernstein & Trahiotis, 1994; Bernstein & Trahiotis, 2002). Some physiological evidence for this phenomenon was observed in the increase in the minimum half-widths of ITD-tuning functions (interpreted as a decrease in the dominant f_m) in response to broadband transposed noise with increasing CF (Chapter 6). This is also likely to be a property of central auditory processing.

In summary, many of the differences between sensitivity to ITDs observed at high-frequencies compared to low-frequencies can be explained by peripheral processing. It is also likely that there is a limit on the f_m at which sensitivity to ITDs can occur which may be a property of the central auditory system.

8.2 The mechanism(s) underlying improved sensitivity to ITDs within transposed tones

Transposed tones were designed to provide high-frequency ANFs with a similar pattern of action potential firing to that observed in low-frequency ANFs in response to pure tones. It was predicted that phase-locking to the f_m of a transposed tone would be tighter than in response to SAM tones of the same f_m. Consistent with this prediction, neural recordings from the IC of the guinea pig had tighter phase-locking (measured by the average vector strength) in response to transposed tones than SAM tones (Figure 3.11). Ideally this comparison would be made from the responses of ANFs and not from neurons in the IC. Tighter phase-locking in response to transposed tones as compared to SAM tones has been observed in ANFs in cats (measured by the vector strength and the height of a shuffled autocorrelogram; Dreyer *et al.*, 2005). The increased modulation depths of ITD-tuning functions in response to transposed as compared to SAM tones are consistent with the model that transposed tones improve sensitivity to ITDs, by increasing phase-locking in high-frequency ANFs (Chapters 3 & 4). The increased modulation depth of ITD-tuning functions is likely to underlie the increased number of neurons and recordings which were found to be sensitive to ITDs.

The increased phase-locking in ANFs in response to transposed tones may be due to an increase in modulation depth, or due to a reduction in the phase of the period of the modulation frequency at which action potentials occur. The mathematical model presented in Chapter 4 suggests that, after peripheral processing, there is a difference between the modulation depths of the "internal" stimuli. The responses of IC neurons to binaurally presented stimuli often occurred over a small portion of the phase of a transposed compared to an SAM tone, especially at the lowest f_m presented. A novel way of investigating the temporal structure of firing rates in ANFs would be to correlate the action potential firing rate, as a function of time, with either the stimulus waveform or with the a model of ANF firing after peripheral processing. Either

Pearson's correlation or the normalised correlation could be used as a measure of correlation and would give different results depending on the unmodulated component (or dc) of the firing rate. It might also be informative to examine the structure of action potential firing in the inputs to coincidence detector neurons i.e. bushy cells in the cochlear nucleus.

Similar to the psychophysical finding reported in Bernstein & Trahiotis (2002), ITD-discrimination thresholds from single neurons were lower in response to transposed tones than SAM tones (Chapter 3). The implication is that neurons in the IC of humans have similar properties and that this forms a physiological substrate of the psychophysical findings. However, the results do not prove that neural JNDs are a relevant (or the only) physiological correlate of psychophysical detection of ITDs. It is not known exactly how neural coding underlies the perception of an ITD, or how ITDs are coded. The lower JNDs in response to transposed tones as compared to SAM tones are merely consistent with the psychophysical findings.

It was of interest that a variety of neural response types were observed, some of which may be useful for detecting ITDs, and some which may not. For example, one neuron (see Section I.4) responded to each "on-period" of transposed tones with 150 Hz f_m, but only to the onset of SAM tones with 150 Hz f_m. This example may represent a single class of neural responses that underlies the enhanced detection of ITDs within transposed tones. It cannot be ruled out that the repetition of information relating to the ITD in response to each "on period" of the transposed tone, from this "class" of neuron might be particularly salient. In general, however, the mean firing rates of neurons were similar in response to SAM and transposed tones (See Figure 4.3).

Transposed tones can be considered to be a series of onsets with different rise times, dependent on their f_m. It is noted that low thresholds for detection of an ITD have previously been obtained using band-pass high-frequency click trains e.g. (Buell & Hafter, 1988). There are similarities between these stimuli and transposed tones

which may be important for detection of ITDs. Clicks trains also contain "off-periods". The "off-periods" may be important for the coincidence detection process itself, alternatively they may simply enable ANFs to follow more precisely the "on period". Experiments altering the length of the "off-period" and the duty cycle of a high-frequency stimulus might provide insights as to the mechanism behind enhanced sensitivity to ITDs in transposed tones. In a transposed tone, the rise time of each "on-period" is dependent on the f_m. This is not the case in a click train. The relationship between sensitivity to ITDs within transposed tones and high-frequency click trains might provide insights into the mechanisms underlying sensitivity to ITDs at high-frequencies (Stecker, 2005).

It is not likely that the enhanced sensitivity to ITDs within transposed tones can be explained simply by considering transposed tones to be a train of onset-ITDs. The results presented in Chapter 5 suggest that onset-ITDs are not an important cue for detection of ITDs, unless ongoing-ITDs are ambiguous. In addition, it is likely that peripheral processing effectively smoothes the transposed waveform such that the "internal" stimulus no longer resembles a train of onsets.

ITD-thresholds are lowest for broadband stimuli (Blauert, 1997). The additional spectral bandwidth of transposed as compared to SAM tones could explain the greater sensitivity to ITDs within transposed tones. The results presented in Chapters 3 and 4 suggest that the improvement in sensitivity to ITDs in a transposed tone can occur within IC neurons with CF equal to the f_c of the stimulus. However, the listener may use neurons with CFs not equal to the f_c. This was not investigated in the experiments described in this thesis. Furthermore, it cannot be discounted that the wider bandwidth of transposed tones allowed for cross-frequency integration which was reflected in the response of a single IC neuron with CF equal to f_c.

In summary, the results presented in the thesis are consistent with the hypothesis that transposed tones improve sensitivity to ITDs by improving the phase-locking to the envelope of a high-frequency stimulus. This, in turn, results in ITD-tuning functions

with greater depths of modulation. Other mechanisms for the improvement in ITD-sensitivity, such as cross-frequency integration, have not been excluded.

8.3 The duplex theory and the physiological relevance of ITDs in the envelope of high-frequency sounds

The duplex theory states that ITDs are used to localise low-frequency sounds, and IIDs are used to localise high-frequency sounds. This appears to hold true for pure tones in the horizontal plane (Stevens & Newman, 1936; Mills, 1960). ITDs and IIDs can be detected within both low- and high-frequency sounds. IIDs can be detected with similar accuracy at low- and at high-frequencies but the IIDs that are naturally present in sounds below 1.5 kHz are small. Low-thresholds for detection of ITDs and high extents of laterality can occur in response to ITDs within both high- and low-frequency complex sounds. However, using band-pass filtered noise (0.5-2 kHz or 4-16 kHz) presented using VAS, it has been shown that ITDs dominate localisation at low-frequencies and IIDs dominate localisation at high-frequencies (Macpherson & Middlebrooks, 2002). This suggests that the duplex theory still holds for the horizontal localisation of complex sounds. The duplex theory does not apply to localisation in elevation and does not explain the resolution of ambiguous ITD and IID cues. In these cases, spectral cues and integration across frequency are required for accurate localisation.

The binaural auditory brainstem pathways which are considered to underlie the duplex theory are also not as dichotomous as the textbook view suggests. High-frequency CFs are over-represented in the LSO and principal neurons are sensitive to IIDs. However, LSO neurons exist that are sensitive to ITDs in the envelope of high-frequency SAM tones (Joris & Yin, 1995) or in low-frequency tones (Tollin & Yin, 2005). More low-CF neurons than high-CF neurons exist in the MSO, but both high- and low-CF neurons in the MSO can be sensitive to ITDs (Yin & Chan, 1990).

Given that sensitivity to ITDs in the envelope of high-frequency sounds can occur, it is of interest to question why ITDs in the envelope of high-frequency sounds are not a more salient cue for sound localisation. The small ITD-discrimination thresholds in response to transposed tones suggest that peripheral processing may explain this; sensitivity to ITDs can only be high as long as appropriate temporal information is present within high-frequency ANFs. In response to broadband noise, Middlebrooks & Green (1990) suggest that the effective "internal" stimulus within a high-frequency ANF may be weakly modulated. Modulation depths of ITD-tuning functions are also likely to be reduced by responses to f_m at which sensitivity to ITDs does not occur. This will tend to occur to a greater extent at the highest CFs; the band-width of ANF filtering increases with increasing CF and, therefore, more f_m above the upper limit at which sensitivity to ITDs occurs will be present in the "internal" stimulus. The weak modulation depth of high-frequency sounds, after peripheral processing, is supported in modelling by MacPherson & Middlebrooks (2002) and in Chapter 7 of this thesis. MacPherson & Middlebrooks (2002) also suggest that ITDs in the envelope of a free-field Gaussian noise may be less useful than when a Gaussian noise is presented over headphones. In the free-field, filtering by the head and pinnae effectively decorrelate the signal at each ear. The lack of comodulation across frequencies could also limit cross-frequency integration, which may otherwise improve detection of ITDs (MacPherson & Middlebrooks, 2002).

In human psychophysical experiments measuring sound localisation, lateralisation or detection of ITDs or IIDs, there is considerable variation between listeners (Blauert, 1997). One listener discussed in MacPherson & Middlebrooks (2002), for example, gave far greater weight to ITDs within a high-frequency band-pass noise than the other listeners. Listeners appear to be able to make use of different cues depending on the cues that are available and their individual preferences. Although ITDs in the envelope of high-frequency sounds are a weak cue for localisation of a noise stimulus, this not does exclude the possibility that they can be used by some listeners in some tasks required.

Evidence that high-frequency envelope ITDs can be useful has been provided by Best *et al.* (2004). In their task, listeners were able to detect the presence, at two different VAS locations, of two concurrent broadband noises using envelope ITD cues. Natural sounds are replete with onsets and offsets and amplitude modulated structures, quite different from the pure tones or Gaussian noise stimuli commonly used in auditory psychophysics. It seems most likely that ITDs within the envelope of high-frequency sounds would be most useful as cues for localisation between 2-4 kHz, where both ITD and IID cues are weak and spectral distortions are not prominent. It would be interesting to investigate how listeners weight ITDs in the envelope of sounds within this frequency range.

It is possible to imagine a selection pressure for development of specialisations for temporal precision in response to low-frequency sound, to enable localisation of frequencies where IIDs are not useful cues. However at higher frequencies directionally-dependent IIDs are available and ITDs become less useful. A selection pressure to follow fast f_m may not exist. However, the high-frequency binaural auditory brainstem does have specialisations for temporal processing; e.g., the aVCN to MNTB synapse, the Calyx of Held, is considered to be the largest and most reliable synapse in the brain. Although some temporal integration in necessary for detection of IIDs, it has been suggested that temporally precise inputs are still required (Tollin, 2003).

Principal neurons in the MSO and LSO neurons have different electrophysiological properties, suggesting that they do play different roles in binaural hearing. LSO neurons are characterised by chopping patterns of action potential firing whereas MSO neurons can faithfully follow their inputs (Tsuchitani & Boudreau, 1969; Smith, 1995). MSO neurons have exceptionally fast membrane time constants and are considered to be specialised for coincidence detection (Svirskis *et al.*, 2004; Magnusson *et al.*, 2005; Scott *et al.*, 2005). Excitatory projections to the IC from the ipsilateral MSO and the contralateral LSO are likely to be segregated (Loftus *et al.*, 2004). However, IC neurons that receive excitatory projections from the ipsilateral

MSO may also receive *inhibition* from low-CF neurons in the contralateral LSO, suggesting that some integration of information from the MSO and LSO does occur (McAlpine *et al.*, 1998; Barnes-Davies *et al.*, 2004; Loftus *et al.*, 2004).

It is interesting to note recent studies that highlight differences in the distributions of ion channels across brainstem auditory nuclei which follow their tonotopic gradient (Brew & Forsythe, 2005; Song *et al.*, 2005). This suggests either different processing occurs across frequency, or it may simply indicate that the same processing requires different properties depending on a neuron's frequency-tuning.

In summary, the duplex theory is probably correct for the localisation of broadband noise in the horizontal plane. Localisation is dominated by IIDs at high-frequencies and ITDs at low-frequencies, but it is incorrect to assume that ITDs cannot be detected at high-frequencies or that IIDs cannot be detected at low-frequencies. The most accurate localisation occurs when directional cues all point to the same location (Gaik, 1993). It is also incorrect to consider that sensitivity to ITDs only originates in the MSO and that sensitivity to IIDs only originates in the LSO; low-frequency LSO neurons may produce functionally-relevant sensitivity to ITDs (Tollin *et al.*, 2005) and IID sensitivity may be generated *de novo* in the IC (Sally & Kelly, 1992). The salience of ITDs in high-frequency sounds is probably limited by the filtering of these sounds by the head and pinnae, and by the peripheral auditory system.

8.4 The potential clinical implications of this work

Current interest in understanding sensitivity to ITDs in the envelope of high-frequency sounds stems not from any finding that envelope ITDs are the most salient cue for sound localisation, but from the trend to implant hearing impaired patients with bilateral cochlear implants. Cochlear implants directly stimulate ANFs (Clark, 2003). The implant electrodes are inserted from the basal end of the cochlea into the scala tympani and are pushed through the spiral of the cochlea towards the apex. The apical end is not reached and therefore low-frequency ANFs are not stimulated.

Sound is recorded through a microphone and cochlear implant processors filter sounds through a band-pass filter to simulate the frequency selectivity of ANFs. The envelope of the waveform in each filter is then sampled. Electrical stimulation consists of a train of pulses which are amplitude modulated according to the envelope derived from the sound waveform (but which are not synchronised interaurally). Therefore, the advantages to be gained from binaural implantation must be derived from the responses of high-frequency ANFs to the envelope of band-pass filtered sounds.

Tests of binaural hearing in bilateral hearing in cochlear implant patients are encouraging: patients prefer listening with two implants and improvements are gained in speech intelligibility in noise and in sound localisation (van Hoesel & Tyler, 2003; Laback *et al.*, 2004; Senn *et al.*, 2005). Most of the improvements in speech intelligibility in noise can be accounted for by the better ear effect (i.e. selectively using the ear with the better signal to noise ratio; Senn *et al.*, 2005). Few patients demonstrate a "binaural squelch" effect, defined as an improvement in speech intelligibility in noise which can be attributed to binaural processing (van Hoesel & Tyler, 2003). JNDs for ITDs, which are only present in the envelope of the electrical trains of impulses, are generally high and are limited to low f_m (van Hoesel & Tyler, 2003; Laback *et al.*, 2004). Increases in localisation performance are attributable to sensitivity to IIDs and not ITDs (van Hoesel & Tyler, 2003).

It is likely that bilateral cochlear implant patients would gain further advantages, particularly in signal detection in noisy environments, if sensitivity to ITDs could be restored. Experiments in this thesis highlight both the potential sensitivity to ITDs that might be obtained using cochlear implants and the limitations in regaining sensitivity to ITDs using cochlear implants. The results obtained using transposed stimuli suggest that high-sensitivity to ITDs in the envelope of high-frequency sounds is possible (Bernstein & Trahiotis, 2002; Bernstein & Trahiotis, 2003; van de & Kohlrausch, 1997; Griffin *et al.*, 2005). However, the greatest sensitivity to ITDs within the envelope of high-frequency sound occurs at relatively low CFs (< 8 kHz)

and for low f_m, optimally around 125 Hz. Therefore, it is likely to be most useful to the cochlear implant patient if ITDs are presented at sites on the cochlear implant electrode that are as apical as possible. It is also likely to be most useful to limit f_m to below ~300 Hz. The presence of higher f_m may reduce the modulation depth of the "internal" stimulus.

Sensitivity to ITDs within the envelope of high-frequencies sounds is limited by peripheral processing which results in a reduced depth of modulation in the firing rates of ANFs. The use of electrical stimulation of ANFs has the potential to by-pass cochlear filtering if the appropriate stimulation is used. This may require the extraction of carrier information at low-frequencies which could be presented by modulation of the amplitude of electrical stimulation, essentially transposing the carrier information to the high-frequency envelope domain. This could be carried out using perhaps just one or two of the most apical electrode channels. Using an electrical stimulus based on the transposed stimulus used by van de Par & Kohlrausch (1997), it has been possible to measure large binaural masking level differences (BMLDs; see Section II.2.8) from bilateral cochlear implant patients (C. Long & R. Carlyon – personal communication).

Although this work is encouraging for users of cochlear implants, cochlear implant patients may never gain hearing performance that is as good as listeners with normal hearing. Reasons for this include damage to or alteration of their central auditory system, the limited number of frequency channels that are currently available and the synchronisation of inputs that occurs across neighbouring ANFs due to receiving the same electrical stimulation. However, improvements in current performance may be possible.

8.5. Conclusions

A neural correlate of the psychophysical improvement in sensitivity to ITDs provided by transposed as compared to SAM tones was identified. ITD-tuning functions recorded from neurons in the IC of the guinea pig have greater depths of modulation, and neural JNDs for ITDs are lower, in response to transposed as compared to SAM tones. Evidence for f_m-dependent ITD-tuning contradicts a model of ITD coding proposed in 1948 by Jeffress, which has dominated physiological investigations and psychophysical models of sensitivity to ITDs for ~50 years. ITD-coding at high-frequencies, similar to ITD-coding at low-frequencies, may use a population of neurons each with broad tuning for ITD.

Sensitivity to ITDs in the envelope of high-frequency sounds is probably restricted both by peripheral processing and also by an upper limitation in the f_m at which sensitivity to ITDs can occur. For these reasons, the physiological relevance of sensitivity to ITDs in the envelope of high-frequency sounds may be limited. The finding that sensitivity to ITDs in the envelope of high-frequency complex sounds can be equivalent to sensitivity to ITDs within low-frequency sounds is encouraging for improving sensitivity to ITDs in cochlear implant users who only receive stimulation of high-CF neurons.

Appendix I – additional data

I.1. Additional data from responses to SAM and transposed tones

I.1.1. ITD-sensitivity assessed from spike rates obtained from different analysis periods

Sixty-nine neurons were investigated using SAM and transposed tones of 500-ms duration with ITDs generated by delaying the entire waveform. In Chapter 3, ITD-sensitivity was assessed using spike rates calculated during a 350 ms analysis window. ITD-sensitivity was also assessed (by calculation of the vector strength and Rayleigh coefficient as described in Chapter 3) using spike rates calculated during a 600-ms window from the onset of the stimulus at one ear to the offset of the stimulus at the other ear. This analysis period included all stimulus-driven spikes. Sensitivity to ITDs was compared using spike rates from each of these time periods. Consistent with the results when data was analysed over the 350 ms time period, when analysed over the 600 ms time period, more recordings and more neurons were sensitive to ITDs within transposed than SAM tones over both analysis windows (Table 4).

More recordings were classified as ITD-sensitive when spikes were analysed over the full (600-ms) analysis window than the middle 350-ms analysis window (Table 4). Recordings were considered ITD-sensitive if the Rayleigh coefficient was >13.815 (see Chapter 3). However, the vector strengths of the recordings were, on average, lower over the 600-ms analysis window as compared to the 350-ms analysis window [Mann-Whitney U test; $p < 0.001$ in response to both SAM (median = 0.16 vs. 0.20) and transposed (median = 0.19 vs. 0.27) tones]. Because the Rayleigh coefficient increases as the vector strength and/or the spike count increases, the likely reason for the increased number of neurons classified as ITD-sensitive when the full 600-ms window was analysed was a result of an increase in the number of spikes counted.

This appears to compensate for the reduced vector strength over the 600-ms analysis window as compared to the 350-ms analysis window.

I.1.2. First spike latency as a function of ITD

Many studies of neural sensitivity to ITDs use either binaural beat stimuli (e.g. Kuwada *et al.*, 1979) or stimuli that are ramped on diotically (e.g. Shackleton *et al.*, 2003). There are few closed-field studies in which ITDs are created by a whole waveform delay, as used in the investigation of neural sensitivity to SAM and transposed tones described in Chapter 3. In the auditory cortex, response latency can provide information as to the location of a sound source (Brugge *et al.* 1996; Furukawa & Middlebrooks, 2002; Stecker *et al.*, 2003; Nelken *et al.*, 2005; Stecker *et al.*, 2005a). Therefore, first spike times in response to SAM and transposed tones at different IPDs were examined. These will be discussed in terms of the first spike latency although the absolute latencies are not presented.

Data Analysis

For 69 neurons (for which spikes were recorded in response to 500 ms stimuli with a whole waveform delay), the times at which the first spikes occurred in response to sound were examined. Particular emphasis was placed on responses to stimuli below 100 Hz f_m in which changes in spike times as a function of ITD were greater and therefore more readily observed. Both raw spike times and "adjusted" spike times were examined. The spike times were adjusted to the onset of sound in the contralateral ear. The first spike times are referred to as the latency of the first spike in response to a sound. However absolute latencies were not calculated.

Neurons were classified into 6 categories according to whether timing of the first spikes in response to SAM and transposed tones, as a function of ITD, were best described as 1) following the onset of sound in the contralateral ear, 2) following the onset of sound at the contralateral and ipsilateral ear, whichever occurred first, 3)

showing evidence of inhibition in response to the onset of sound at the ipsilateral ear indicated by a reduction in spontaneous rate or 4) by showing an increase in latency in response when stimulation at the ipsilateral ear led stimulation at the contralateral ear, 5) having a trough in firing rate and an absence of a first spike at some ITDs or 6) whether the first spike latency did not fit into any of the first 5 categories.

Results and discussion

An example of neural responses which followed the onset of sound at the contralateral ear is shown in Figure I.1. This Figure shows raster plots of spike times in response to SAM and transposed tones at f_m = 10 Hz, obtained from a neuron that was insensitive to ITDs. In the top two panels, (A and B), the raw spike times are displayed, whereas in the bottom two panels (C and D) the spike times displayed were adjusted to account for the latency of sound arrival at the contralateral ear (see *Data analysis* above). When adjusted for differences in the latency of the onset of sound at the contralateral ear, the first spike latency was constant as a function of the IPD presented. The first spike latency followed the onset of sound in the contralateral ear for over half of the neurons that were insensitive to ITDs (Table 5; under the heading "contra").

Figure I.2 shows spike times recorded from two neurons for which the first spike latency followed the first onset of sound, whether this occurred at either ear. Spike times displayed were adjusted according to the onset of the contralateral ear. When the IPD was negative, the first spike latencies of both neurons followed the onset of sound in the contralateral ear. For neuron 25606, when the IPD was negative, the latency of the first spike was shorter than predicted from the onset of sound at the contralateral ear. The latency of the first spikes recorded from this neuron followed the onset of sound at the ipsilateral ear (Figures I.2A & B). This neuron was insensitive to ITDs, whereas neuron 25902 (Figure I.2C & D) was sensitive to ITDs. For neuron 25902, at negative IPDs and particularly at IPDs close to -0.5 cycles, first spike latency also decreased, suggesting that this neuron was excited by sound

presented to the ipsilateral ear. This neuron was ITD-sensitive with a minimum in firing rate at IPDs around 0.0 cycles (at modulation rates from 14 to 42 Hz in response to transposed tones) and in response to *pure tones* at the ipsilateral ear (not shown), there was evidence for inhibition of spike firing in response to the offset of the tone. Taken together, these results suggest that neuron 25902 responded with a mixture of inhibitory and excitatory interactions in response to sound at the ipsilateral ear. Table 5 shows the total number of neurons for which first spike latency followed the onset of sound in the contralateral ear at positive IPDs but decreased at negative IPDs, consistent with excitation in response to the ipsilateral ear ("contra and ipsi"). There was no neuron for which first spike latency followed the ipsilateral ear exclusively. This is consistent with the observation that all neurons tested were excited by tones presented to the contralateral ear alone (see Chapter 3).

The number of neurons for which there was evidence for inhibition in response to sound at the ipsilateral ear is shown in Table 5 under two headings: "inhibition shown by reduced spontaneous" and "inhibition shown by increased latency". Two neurons showing a reduced spontaneous rate in response to ipsilateral stimulation are shown in Figure I.3. The first spikes shown in the raster plots, particularly at negative IPDs, are not evoked spikes but represent the spontaneous activity of these neurons. It is possible to observe the effect of inhibition is response to the ipsilateral ear in the reduction in spontaneous rate at negative IPDs. The phase-locking of neuron 25906 (Figure I.3A&B) is modulated as a function of IPD, despite the fact that the firing rate of this neuron was not sensitive to ITDs. The phase-locking of neuron 25808 (Figure I.3C&D) was not modulated as a function of ITD. However, this neuron was sensitive to ITDs at higher modulation rates with a minimum firing rate around ITD = 0 μs, consistent with interaction between excitatory and inhibitory responses from the contralateral and ipsilateral ears.

The first spike latency of some neurons (see Table 5) increased at negative ITDs. It is likely that the increase in latency was due to inhibition in response to sound at the

ipsilateral ear. An example of this increase in latency with IPD is shown in Figure I.4 for a neuron that was sensitive to ITDs at higher modulation rates than the 10 Hz f_m shown. All but one of the neurons in this category, for which monaural responses were obtained, demonstrated inhibition in response to ipsilateral tones. The exception produced an onset response to a tone presented at the ipsilateral ear.

Onset latency, as a function of ITD, was often more complex for ITD-sensitive neurons than for neurons that were insensitive to ITDs. For nine ITD-sensitive neurons, the first spike was "absent" close to 0-μs ITD, where there was also a minimum in the firing rate of the neuron. An example of this is shown in Figure 9.1.5B, in response to transposed tones with 210 Hz f_m. The responses of these neurons commonly showed evidence for inhibition in response to sound presented at the ipsilateral ear in the form of reduced latencies at negative IPDs (seen particularly at f_m = 10 Hz). 5/9 of these neurons showed a reduction in spontaneous rate, and none were excited, in response to tones presented at the ipsilateral ear.

Figures I.5C & D show raster plots of an example response from the final classification of onset latency. In this group, onset latencies were modulated with ITD but were not easily described by one of the previous categories. This group contains neurons that were both considered to have maximal responses close to 0-μs ITD and those that have minimum response close to 0-μs ITD.

In summary, information regarding the ITD in the stimulus was present in the first spike times of many of the neural responses obtained. This was most obvious at low f_m where the ITDs presented were largest. The amount of information in these spike times was not quantified. The examples provided in Figures I.1.14 & 5 demonstrate differences in the first spike latencies in response to SAM and transposed tones. However, no consistent differences were noted in response to SAM and transposed tones.

I.1.3. Comparison of firing rates in response to SAM and transposed tones with + 0.5 cycles and -0.5 cycles IPD

Spike times were obtained in response to SAM and transposed tones at ± 0.5 cycles IPD (with reference to the cycle of the f_m), for the same neurons. In terms of an ongoing-ITD, + 0.5 cycles IPD and – 0.5 cycles are equivalent; if stimuli were presented by ramping on the sounds diotically in each ear, stimuli presented at + 0.5 and – 0.5 cycles would be identical. However, in this investigation, the sounds had on onset-ITDs which were different depending on whether the IPD was + 0.5 cycles or - 0.5 cycles. In Chapter 3, the firing rates at ± 0.5 cycles were pooled and the average firing rate, within the middle ~350 ms of the response was used to assess sensitivity to ITDs. In this Section, the firing rates at ± 0.5 cycles are compared to assess whether the difference in onset-ITD influences the firing rate to the ongoing-ITD which, in the case of ± 0.5 cycles, might be predicted to be identical.

The histograms in Figure I.6 depict the distribution of differences in firing rates at ± 0.5 cycles (firing rate at - 0.5 cycles minus firing rate at + 0.5 cycles) for all recordings in response to SAM (A) and Transposed (B) tones. There was no statistical evidence that firing rates were dependent on whether the IPD was ± 0.5 cycles ($p = 0.41$, a 4-factor ANOVA testing for main effects of the f_m, neuron, stimulus type and whether the recording was at ± 0.5 cycles).

Histograms of the difference in firing rate in response to ± 0.5 cycles ITD for recordings which were classified as sensitive to ITDs are shown in Figure I.6C & D. ITD-sensitive recordings were examined separately from ITD-insensitive recordings. The difference in onset-ITD had little effect on the firing rate measured in response to the ongoing portion of the stimulus ($p = 0.78$, ANOVA testing the main effects of the neuron, f_m, stimulus type and whether the recording was at + 0.5 cycles or – 0.5 cycles). Differences between neural responses to ± 0.5 cycles ITD are predominantly in the latency of the onset of the sound (see Section I.1.2).

I.1.4. Rate modulation transfer functions

Firing rate as a function of the modulation rate in response to SAM and transposed tones presented was compared.

Data Analysis

Firing rates in response to SAM and transposed tones, over a range of f_m [the rate modulation transfer function (rMTF)], were classified according to whether they were dominated by a peak in firing rate ("band-pass"; Figure I.7A), whether they were dominated by a trough in firing rate over a range of f_m ("band-reject"; Figure I.7B), whether they displayed reduced firing rates at low f_m ("high-pass"; Figure I.7C), or whether they displayed higher firing rates at low f_m ("low-pass"; Figure I.7D). Neurons could also be classified as "not modulated" in their responses.

For this classification, only responses to ITDs of 0 μs were examined from neurons to which transposed and SAM tones were presented and at f_m between 10 and at least 300 Hz. Neurons were classified as "not modulated" if their minimum firing rate (Figure I.7; blue squares) was greater than 70% of the maximum firing rate (Figure I.7; red asterisks). Modulated rMTFs were further classified as follows. The f_m at which the firing rate was highest was first identified ($f_m max$). Neurons were classified as "band-pass" if f_ms both higher and lower than $f_m max$ showed firing rates lower than 70% of that at $f_m max$ (Figure I.7A, black circles). The f_m at the lowest firing rate was then identified ($f_m min$). If f_ms both higher and lower than $f_m min$ showed firing rates greater than 50% [(maximum-minimum) + minimum] (black circles), the response was classified as "band-reject" (Figure I.7B). If the only $f_m s$, with less than 70% maximum firing rate (black circles) were below $f_m max$ the response was classified as "high-pass" (Figure I.7C). If the only values of f_m with less than 70 % maximum firing rate (black circles) were higher than $f_m max$ the neuron was classified as "low-pass" (Figure I.7D).

Results

Responses of 71 neurons were examined for which SAM and transposed tones were presented with IPD = 0 cycles and at f_m between 10 and at least 300 Hz. The firing rates of these neurons, at 0-cycles ITD, describe their rate modulation transfer function (rMTF). Firing rates were considered modulated as a function of f_m if the minimum firing rate was <= 70 % of the maximum firing rate. Modulated rMTFs therefore comprised firing rates varying by at least 30 % maximum firing rate over the range of f_ms tested. In response to SAM tones, 12 neurons did not meet this criterion and were classified as "not modulated" while, in response to transposed tones, only one neuron did not meet the criterion ($p<0.01$, test for independence of paired proportions, see Table 6). The result that more neurons had modulated rMTFs in response to transposed tones compared to SAM tones was robust; the criterion that was chosen for the classification could vary between 10 and 90 % maximum firing rate. In general, rMTFs were more modulated in response to transposed compared to SAM tones (normalised firing rate changes [measured as the maximum-minimum/maximum firing rates of the rMTF] were greater in response to transposed than SAM tones; paired t test; $p < 0.01$).

rMTFs were further classified according to the scheme described in *Data analysis* above. Note that a comprehensive range of f_ms was not examined for all neurons, neither were f_ms sampled equally for each neuron. The rMTF characterization was, therefore, not definitive. The number of neurons classified as "band pass", "band reject", "low-pass" and "high-pass" is shown in Figure I.8 according to whether the rMTF was in response to SAM or transposed tones. The rMTF classifications of each of the example neurons in Figures 3.1-5 are displayed in Table 7.

rMTFs appeared qualitatively similar in response to SAM and transposed tones (as can be seen in Figures 3.1-5). In agreement with this, 33 neurons were assigned the same classification in response to transposed and SAM tones. Of those 38 neurons

that were assigned different classifications in response to transposed and SAM tones, 11 were classified as not modulated in response to SAM tones but were modulated in response to transposed tones (e.g. Figure 3.5C). Nine neurons classified as "band-pass" in response to transposed tones were classified as "low-pass" in response to SAM tones (e.g. Figures 3.1 & 3.5B). The differences in rMTF in response to SAM and transposed tones were often consistent with the increased sensitivity to ITDs in transposed tones (e.g. Figures 3.5A-F & Figures 3.1&2). However, differences in rMTF in response to SAM and transposed tones also occurred in responses from neurons that were not sensitive to ITDs (e.g. Figure 3.5G). Overall, rMTFs in response to transposed tones and from ITD-sensitive neurons were most likely to be classified as either band-pass or band-reject, consistent with the larger modulation in rMTFs in response to transposed tones and with the influence of the greater sensitivity to ITDs in response to transposed tones (which occurred at specific f_m).

rMTFs were only classified at 0-cycles IPD. It was clear that, if the rMTF had been examined at another value of IPD, e.g. 0.25 cycles, the modulation in firing rate of ITD-sensitive neurons as a function of IPD, would have altered the rMTF classification (e.g. Figures 3.1 & 2 and Figures 3.5A, B, D & E). In other cases (e.g. Figure 3.5C & F) it is unlikely that sensitivity to ITDs, *per se*, generated the rMTF classification. It is envisaged that if rMTFs were examined at a constant ITD, for some but not all neurons, the rMTF classification would depend on the ITD examined.

I.1.5. Calculating neural JNDs for ITD from spike rates obtained from different analysis periods

In Section 3.3.3, neural JNDs for detection of an ITD from 0-μs ITD were presented. Neural JNDs were calculated from neural firing rates obtained within a 350-ms response window. Neural firing rates were also obtained from a 600-ms response window, incorporating both the onset and offset responses to SAM and transposed tones (see Section 2.1.7). Figure I.9 displays neural JNDs calculated from the 600-ms

response window plotted as a function of neural JNDs calculated from the 350-ms response window. JNDs were generally lower in response to transposed tones irrespective of which response window was examined, and the lowest thresholds were obtained from responses obtained over the 600-ms response window. Two data points in response to SAM tones represent JNDs obtained over the 600-ms window but not the 350-ms window (marked by arrows in Figure I.9). The data point representing the lowest JND in response to SAM tones suggests that, by including the onset (and offset) response to SAM tones, neural JNDs can be very small (at least in the case of one neuron) in response to SAM tones. Raster diagrams of spike times from this neuron are shown in Figure I.10. The response to sound onset was well modulated by ITDs in response to SAM and transposed tones whereas the neuron only responded to the ongoing stimulus in response to transposed tones. This neuron responded to pure tones, of frequency equal to its CF, with an onset response. It was not the case that all neurons fired more strongly to the ongoing component of transposed tones than SAM tones (see Section 4.3). However, in this case, sensitivity to ITDs in transposed tones was greater than sensitivity to ITDs in SAM tones because transposed tones at 150 Hz f_m elicited spikes throughout the duration of the stimulus whereas SAM tones at 150 Hz f_m did not.

I.2.Additional data from responses to transposed noise

I.2.1. Monaural responses to broadband transposed noise

Firing rates of 22 neurons were compared in response to broadband transposed noise that was presented monaurally to the each ear and also presented diotically. The firing rate in response to a stimulus presented to the ipsilateral ear alone was always lower than the response to a stimulus presented to the contralateral ear alone or to both ears. For most neurons (18/22), firing rates were greater to contralateral than to diotic stimulation. Of 21 neurons for which monaural stimuli were presented and that were also presented with transposed noise stimuli with different ITDs, 4 were sensitive to

ITDs. These 4 neurons responded with a higher firing rate to contralateral sound than to sound presented diotically although 2 were peak-types (and 2 were trough-types). That high-CF neurons in the IC are predominantly driven by contralateral excitation has been noted previously (e.g. Batra *et al.*, 1993; Chapter 3).

I.2.2. The effect of stimulus intensity on ITD-tuning

ITD-tuning functions from 4 neurons were obtained in response to transposed noise at different intensities and these are displayed in Figure I.11. For one neuron (Figure I.11D), an ITD-tuning function obtained with a carrier frequency above CF was also included. In this case, the cut-off of the low-frequency modulating noise was equal to $f_c/5$. The overall intensity of the stimulus could alter the firing rate of neurons to a greater extent than different ITDs at any single intensity. However, the shapes of ITD-tuning functions obtained at different intensities from the same neurons were similar. In the analyses presented in Chapter 6, the ITD-tuning function with the greatest modulation depth (in sp/s) for each neuron where ITD-functions were obtained at different intensities was shown.

Yin *et al.* (1986) reported that ITD-tuning functions in response to wide-band noise were relatively invariant with the intensity of the stimulus. Although ITD-tuning functions in response to broadband transposed noise had broadly similar shapes in response to different stimulus intensities, the mean firing rate was strongly affected by the intensity of the stimulus. Phase-locking in response to transposed and SAM tones in ANFs has been reported to be more strongly affected by intensity than phase-locking in response to low-frequency pure tones (Dreyer *et al.*, 2005).

I.2.3. Firing rates in response to correlated and uncorrelated broadband transposed noise

Responses to correlated and uncorrelated broadband transposed noise were initially obtained to establish an appropriate intensity at which to present transposed noise

stimuli (see Section 2.1.7). Firing rates from 56 neurons in response to correlated and uncorrelated noise are compared in Figure I.12. Firing rates from 13 neurons sensitive to ITDs are shown in Figure I.12A according to whether they were classified as peak-type (black) or trough-type (blue). Firing rates from 43 neurons insensitive to ITDs are shown in Figure I.12B. Consistent with peak-type neurons responding with a maximum firing rate close to 0-μs ITD, firing rates of peak-type neurons tended to be higher in response to correlated as compared to uncorrelated transposed noise. Consistent with trough-type neurons responding with a trough in firing rate close to 0-μs ITD, firing rates of trough-type neurons tended to be lower in response to correlated as compared to uncorrelated transposed noise. Firing rates in response to correlated minus uncorrelated broadband transposed noise were higher for peak-type neurons than trough-type neurons (Mann-Whitney U test, $p < 0.0001$). Overall the absolute difference between firing rates of ITD-sensitive neurons in response to correlated and uncorrelated transposed noise was small (median = 2.2 sp/s, maximum = 30 sp/s). The differences in firing rates of neurons that were not sensitive to ITDs in response to correlated and uncorrelated noise were slightly lower than this (median = 1.6 sp/s, maximum = 17 sp/s, Mann-Whitney U test, $p = 0.027$).

I.2.4. ITD-tuning in response to transposed noise with different bandwidths

Responses from 13 neurons were obtained in response to transposed noise stimuli with the low-pass cut-off frequency of the modulating noise at a frequency other than CF/5 (see Table 8A). In this Section, the effect of changing the bandwidth of the modulating noise on the ITD-tuning functions obtained will be described.

Figure I.13 shows ITD-tuning functions from 4 neurons that were sensitive to ITDs in response to broadband transposed noise (black dashed line) as well as transposed noise with a lower cut-off in envelope frequency (red line). Firing rates were filtered with a 3-point box car and normalised between 0 and 1 to emphasise the tuning to

ITDs. The spectrum of the modulating noise was flat such that although the attenuation of the stimulus was constant, the intensity of the stimulus decreased when the band-width of the modulating noise decreased. In one case in Figure I.13 the attenuation of the stimulus was decreased when the band-width of the stimulus envelope was decreased (Figure 6.6.1D). In this case, the attenuation is noted in the Figure key. ITD-tuning functions in Figures I.13A-C were only slightly wider in response to the transposed noise with the lower cut-off frequency (red). The lower cut-off frequency was high in comparison range of f_m at which sensitivity to ITDs in the envelope of high-frequency stimuli was found to occur (see Chapter 3). In Figure I.13D, the ITD-tuning function in response to transposed noise with 100 Hz low-pass cut-off was wider than the ITD-tuning function in response to a cut-off of CF/5 (636 Hz; although note that these ITD-tuning functions were obtained at different values of attenuation and therefore the flat spectrum of he stimulus was at a different level).

ITD-tuning functions obtained from one neuron from which the most ITD-tuning functions were obtained are shown in Figure I.14. Responses of this neuron were obtained to presentation of transposed noise stimuli that were both low-pass (Figure I.14A) and band-pass (Figure I.14B, C & D) for f_m. The response to the broadband transposed noise (black dotted line) is repeated in each of Figures I.14A-D to assist comparisons between these Figures. The ITD-tuning function in response to broadband transposed noise was classified as peak-type but was characterised by both a peak and a trough in firing rate. When the low-pass cut-off of the modulating noise was decreased to 500 or 200 Hz, the ITD-tuning functions appeared similar (Figure I.14A). However, when the low-pass cut-off was 100 Hz, the ITD-tuning function broadened and was characterised predominantly by a trough in firing rate (Figure I.14A). A trough in firing rate also characterised the ITD-tuning functions in response to transposed noise that was band-pass for f_m and centred on 125 Hz (Figure I.14B). However, when the band-pass transposed noise was centred on 50 Hz f_m, both a peak and trough in firing rate were apparent (Figure I.14C).

It was predicted that ITD-tuning functions in response to a stimulus that was band-pass for f_m would be quasi-periodic, as seen in the responses of low-CF neurons to broadband Gaussian noise. This could be observed in the firing rates in response to a band-pass transposed noise centred on 250 or 400 Hz f_m (Figure I.14D). ITD-tuning functions in response to these stimuli were weakly modulated and not significantly sensitive to ITDs. However, a clear periodicity can be observed in the ITD-tuning functions. This periodicity may not have been observed in Figures I.14B & C because the ITD-tuning functions were not sampled at large enough ITDs. Note that the periodicity in response to transposed noise with f_m centred on 400 Hz (shown in red in Figure I.14D) suggests that this neurons was capable of (at least weak) sensitivity to ITDs at f_m around 400 Hz.

The peak-ITDs and trough-ITDs of all ITD-tuning functions that were ITD-sensitive are shown as a function of their half-width in Figures I.15A & B. (The half-widths of all ITD-tuning functions are shown in Figure 6.7 as a function of the cut-off of the low-pass modulating noise.) At wide half-widths, the estimate of the peak or trough-ITDs was less reliable due to broad peaks and troughs. The linear regression between peak-ITD and half-width was still significant ($p = 0.0041$) but only accounted for 26% of the variance.

Analysis Period	Total	Sensitive to ITDs in trans. Tones only	sensitive to ITDs in trans. and SAM tones	sensitive to ITDs in SAM Tones only	insensitive to ITDs
Number of neurons					
600 ms	69	17	25	3	24
350 ms	82	23	23	1	35
Number of recordings					
600 ms	809	161	84	30	534
350 ms	952	172	57	24	706

Table 4. The number of neurons and recordings sensitive to ITDs, analyzed over a 600 ms and 350 ms period.

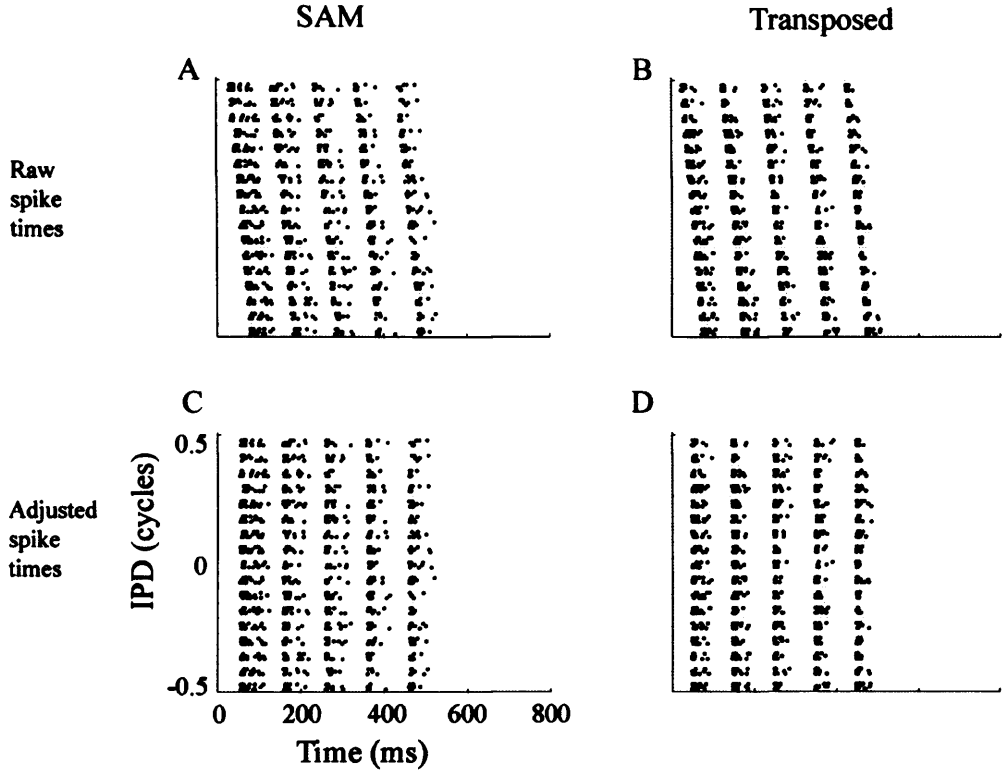

Figure I.1. Raster diagrams of spike times in response to SAM (A & C) and transposed (B & D) tones with 10 Hz f_m. Raw spike times are displayed in A&B whereas spike times adjusted for the onset of sound in the contralateral ear are shown in C&D. Neuron 29305, CF = 2.9 kHz.

	Number of neurons	
Classification of first spike latency	ITD sensitive	ITD in-sensitive
1. Contra	3	16
2. Contra and ipsi	6	4
3. Inhibition observed by reduced spontaneous	8	6
4. Inhibition observed by increased latency	6	3
5. First spike absent in trough	9	0
6. Complex	8	0
Total	40	29

Table 5. The number of neurons in each of 6 classifications of first spike latency. Classifications are described in Section 2.1.7.4 discussed in Section I.1.2.

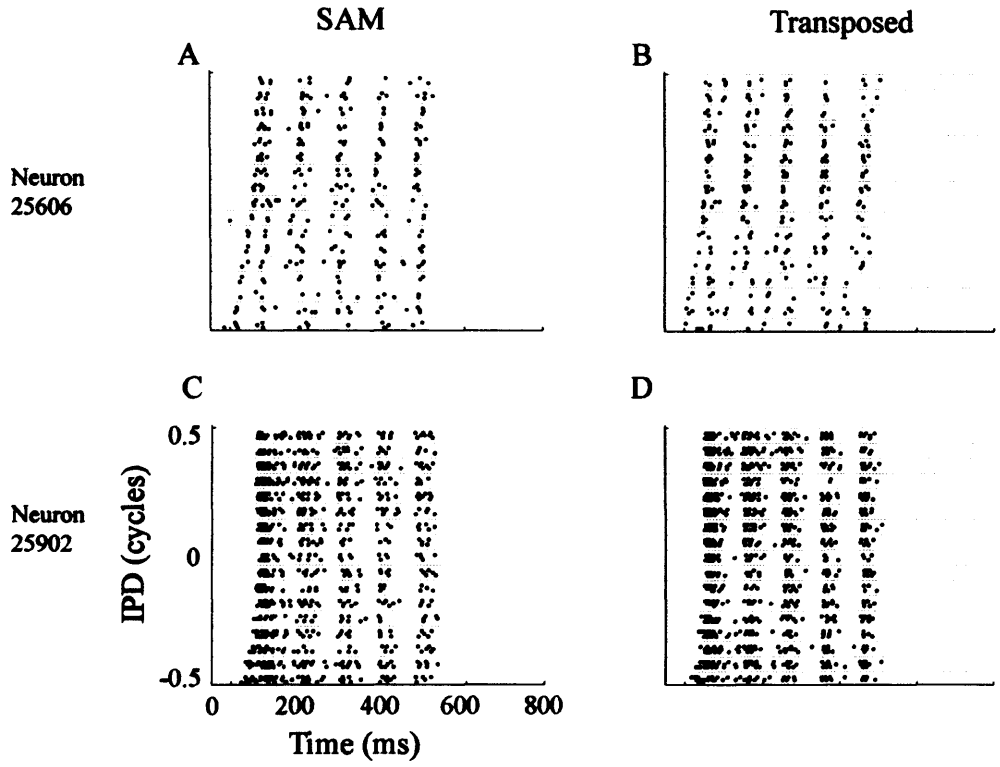

Figure I.2. Raster diagrams of spike times adjusted for the onset of sound in the contralateral ear in response to SAM (A & C) and transposed (B & D) tones with 10 Hz f_m. A&B) Neuron 25606, CF = 6.5 kHz C&D) Neuron 25902, CF = 3.9 kHz.

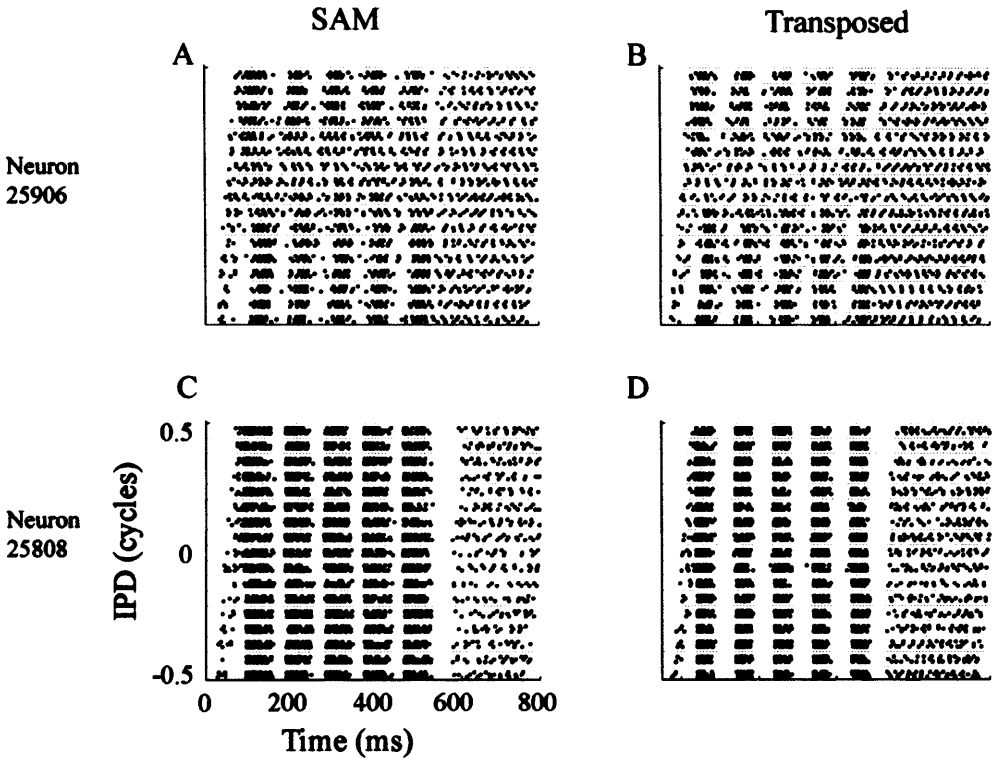

Figure I.3. Raster diagrams of spike times adjusted to the onset of sound in the contralateral ear in response to SAM (A & C) and transposed (B & D) tones with 10 Hz f_m. A&B) Neuron 25906, CF = 6.8 kHz C&D) Neuron 25808, CF = 4.6 kHz.

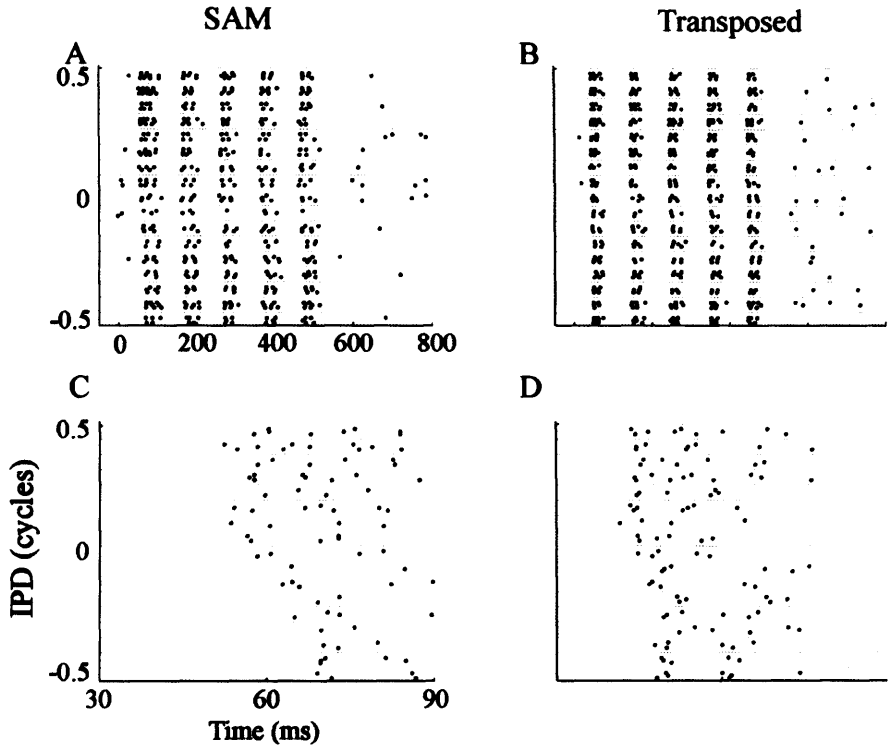

Figure I.4. Raster diagrams of spike times adjusted to the onset of sound in the contralateral ear in response to SAM (A & C) and transposed (B & D) tones with 10 Hz f_m. Neuron 23720, CF = 8.1 kHz

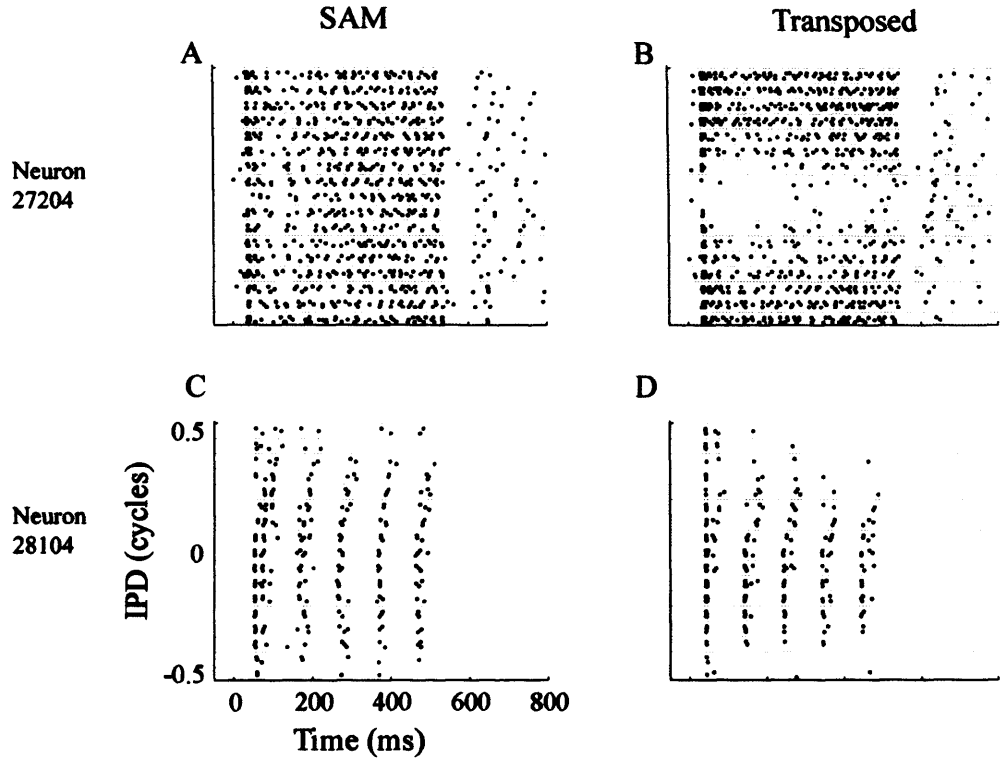

Figure I.5. Raster diagrams of spike times adjusted to the onset of sound in the contralateral ear in response to SAM (A & C) and transposed (B & D) tones with 210 Hz (A&B) or 10 Hz (C&D) f_m. A&B) Neuron 27204, CF = 2.2 kHz C&D) Neuron 28104, CF = 3.4 kHz.

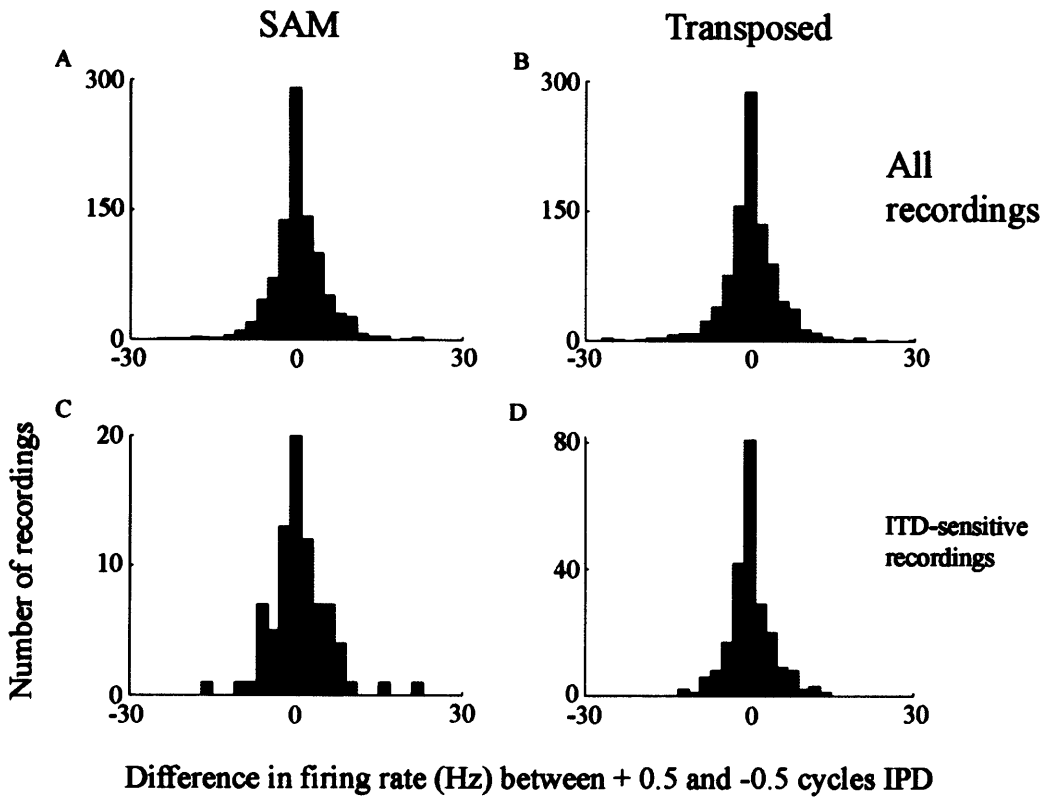

Figure I.6. The difference in firing rates between -0.5 and +0.5 cycles of all recordings in response to SAM (A) and transposed tones (B). The difference in firing rates between -0.5 and +0.5 cycles of all recordings that were classified as ITD-sensitive in response to SAM (C) and transposed tones (D).

		Responses to Transposed tones		
		not modulated	modulated	Total
Responses to SAM tones	not modulated	1	11	12
	modulated	0	59	59
	Total	1	70	71

Table 6. The number of neurons which had rMTFs that were classified as modulated and not modulated in response to SAM and transposed tones.

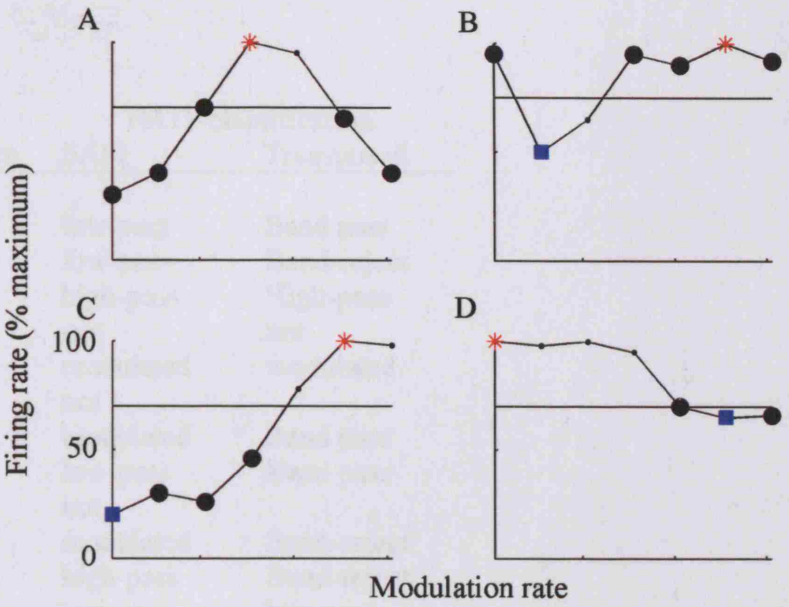

Figure I.7. A diagram to illustrate how rMTFs were classified as A) Band-pass, B) Band-reject, C) High-pass and D) Low-pass. The rMTFs shown are not neural responses but were constructed to illustrate the rMTF classification. Blue squared mark the minimum firing rate of each rMTF and red asterisks mark the maximum firing rate of each rMTF. Black horizontal lines mark 70% of the maximum firing rate (A, C & D) or half-way between the maximum and minimum firing rate (B).

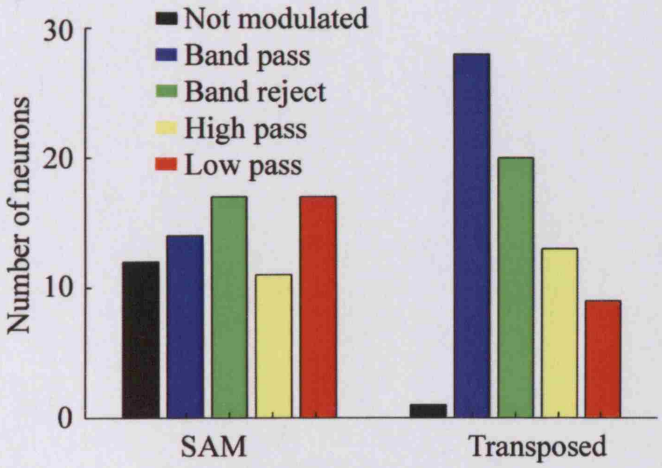

Figure I.8. The classification of neurons' rMTFs in response to SAM and transposed tones. See section 9.3 and 2.1.7.

	rMTF classification	
Figure	SAM	Transposed
3.1	low-pass	Band pass
3.2	low-pass	Band-reject
3.3	high-pass	High-pass
	not	not
3.4	modulated	modulated
	not	
3.5A	modulated	Band pass
3.5B	low-pass	Band pass
	not	
3.5C	modulated	Band-reject
3.5D	high-pass	Band-reject
3.5E	low-pass	low-pass
3.5F	band-reject	Band-reject
3.5G	high-pass	Band-reject
3.5H	high-pass	High-pass
	not	
3.5I	modulated	High-pass

Table 7. The rMTF classification of the neural responses shown in Figures 3.1-5.

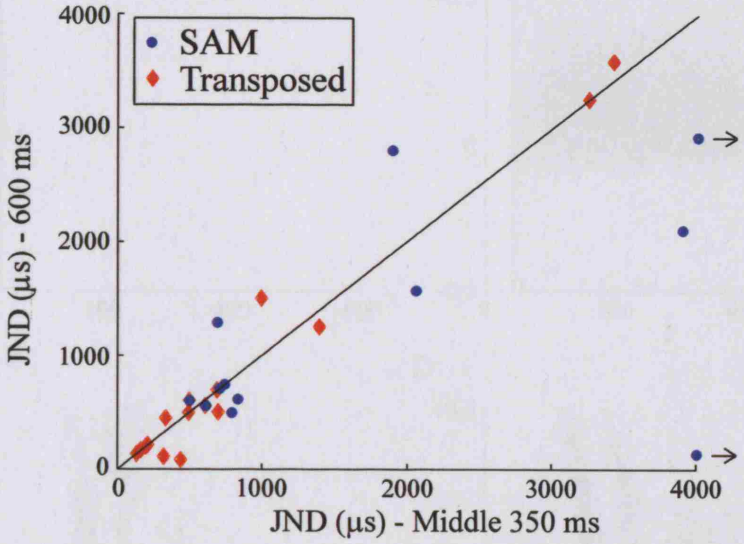

Figure I.9. Neural JNDs calculated from firing rates during a 600-ms or a 350-ms window (see methods) in response to SAM (blue) and transposed (red) tones. Arrows mark data in response to SAM tones for which there was no threshold of 4 ms or below.

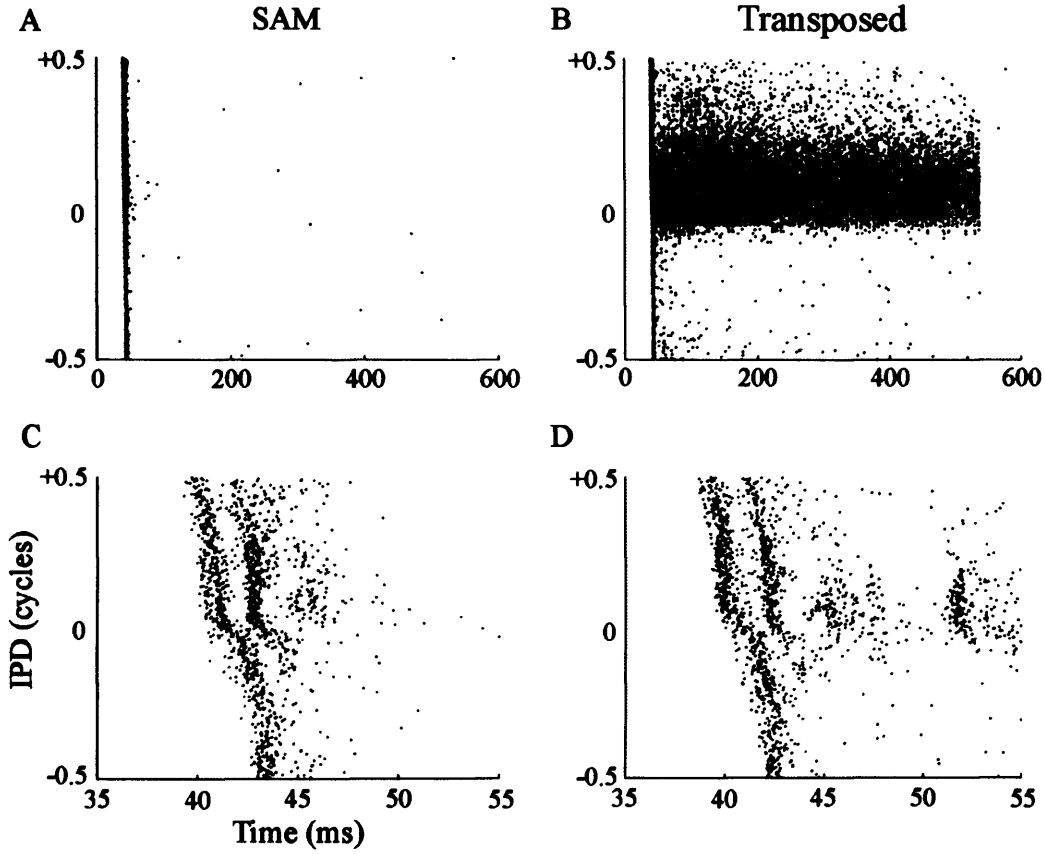

Figure I.10. Raster plots of spike times of neuron 25806, CF = 7.8 kHz, in response to SAM (A&C) and transposed (B&D) tones at 150 Hz f_m. Spike times over 600 ms are shown in Figures A&B whereas spike times in response to the sound onset are shown in Figures C&D.

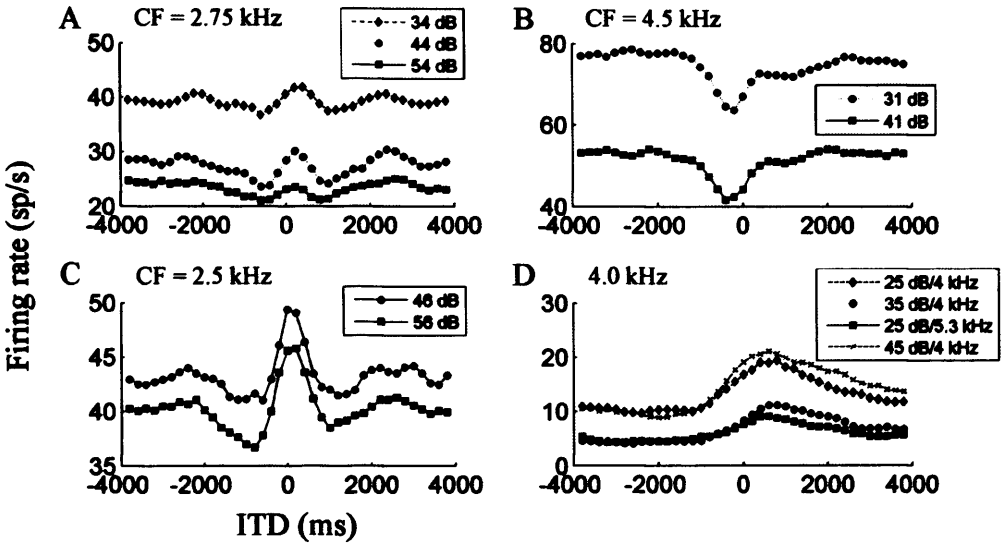

Figure I.11. The mean firing rates (filtered with a 3 point box car) from 4 neurons from which more than one ITD-tuning curve was recorded. Neuron numbers: A) 28803, B) 28806, C) 28805, D) 34311. The ITD-tuning function in B was classified as trough-type and the others three were classified as peak-type. The dB attenuation at which the stimulus was presented is shown in the legend. One response to a transposed noise stimulus with f_c not equal to CF is shown in D.

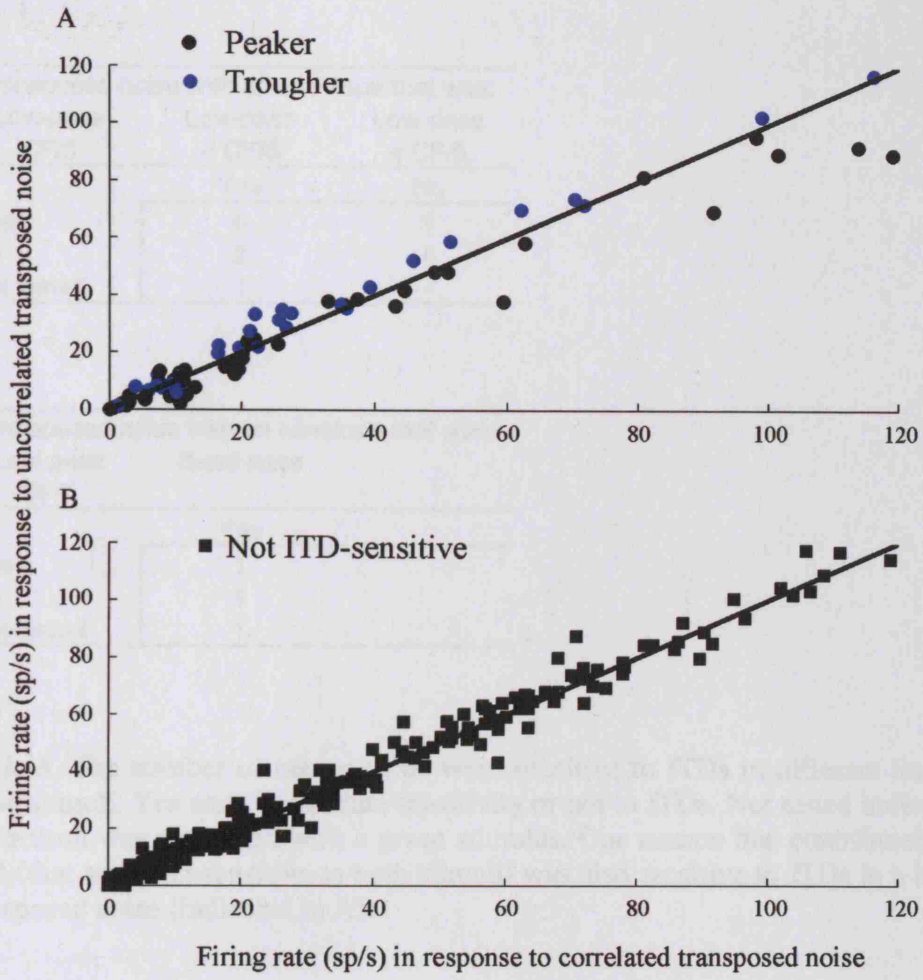

Figure I.12. Firing rates in response to correlated and uncorrelated broadband transposed noise.

A

Transposed noise with an envelope that was:		
Low-pass CF/5	Low-pass < CF/5	Low-pass < CF/5
	Yes	No
Yes	6	1
No	2	4
Not tested	1	-

B

Transposed noise with an envelope that was:	
Low-pass CF/5	Band-pass
	Yes
Yes	1
No	1
Not tested	1

Table 8. The number of neurons that were sensitive to ITDs in different transposed noise stimuli. Yes and No indicate sensitivity or not to ITDs. Not tested indicates that the neuron was not tested with a given stimulus. One neuron that contributed to data in B (that was ITD-sensitive to both stimuli) was also sensitive to ITDs in a low-pass transposed noise (indicated in A).

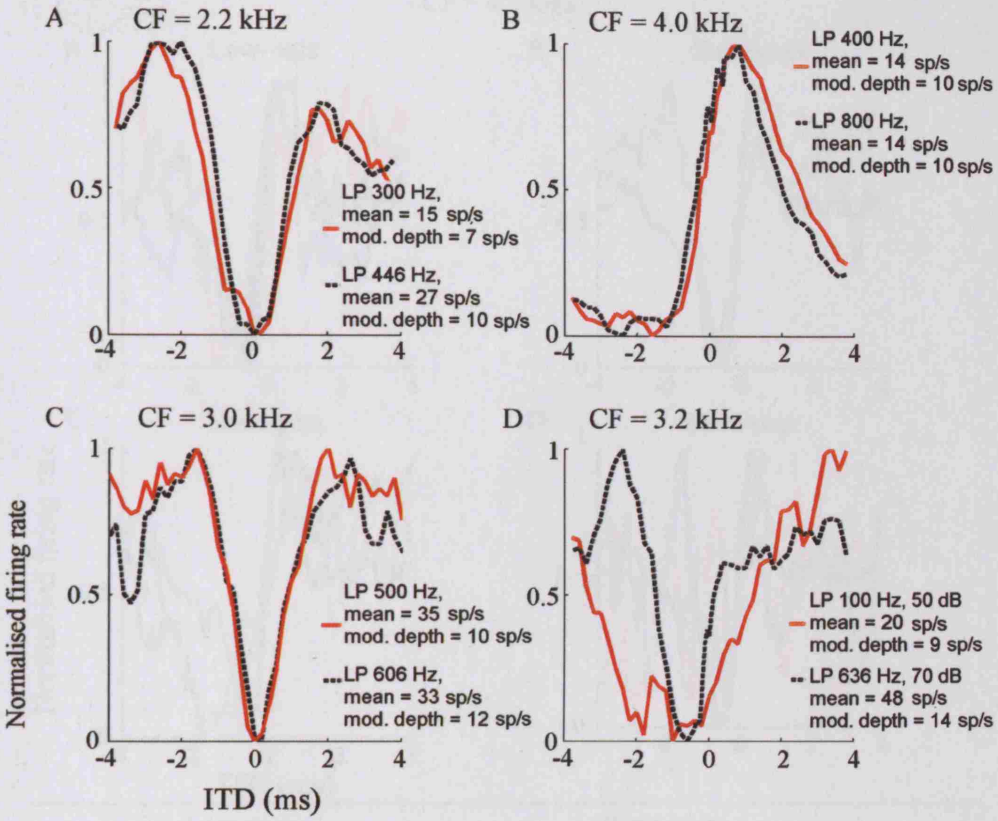

Figure I.13. Normalised ITD-tuning curves generated from the responses of 4 neurons (A-D), in response to transposed noise with the low-pass cut-off of the modulating noise equal to CF/5 (dotted black line) and with one other low-pass cut-off (red line). Keys display the low-pass (LP) cut-off frequency of the stimulus envelope and the mean firing rate and modulation depth of the ITD-tuning curves, before normalisation. A) neuron 27204, attenuation = 85 dB. B) neuron 34311, attenuation = dB = 35 dB. C) neuron 42907, attenuation = 60 dB. D) neuron 34108, attenuation marked in key.

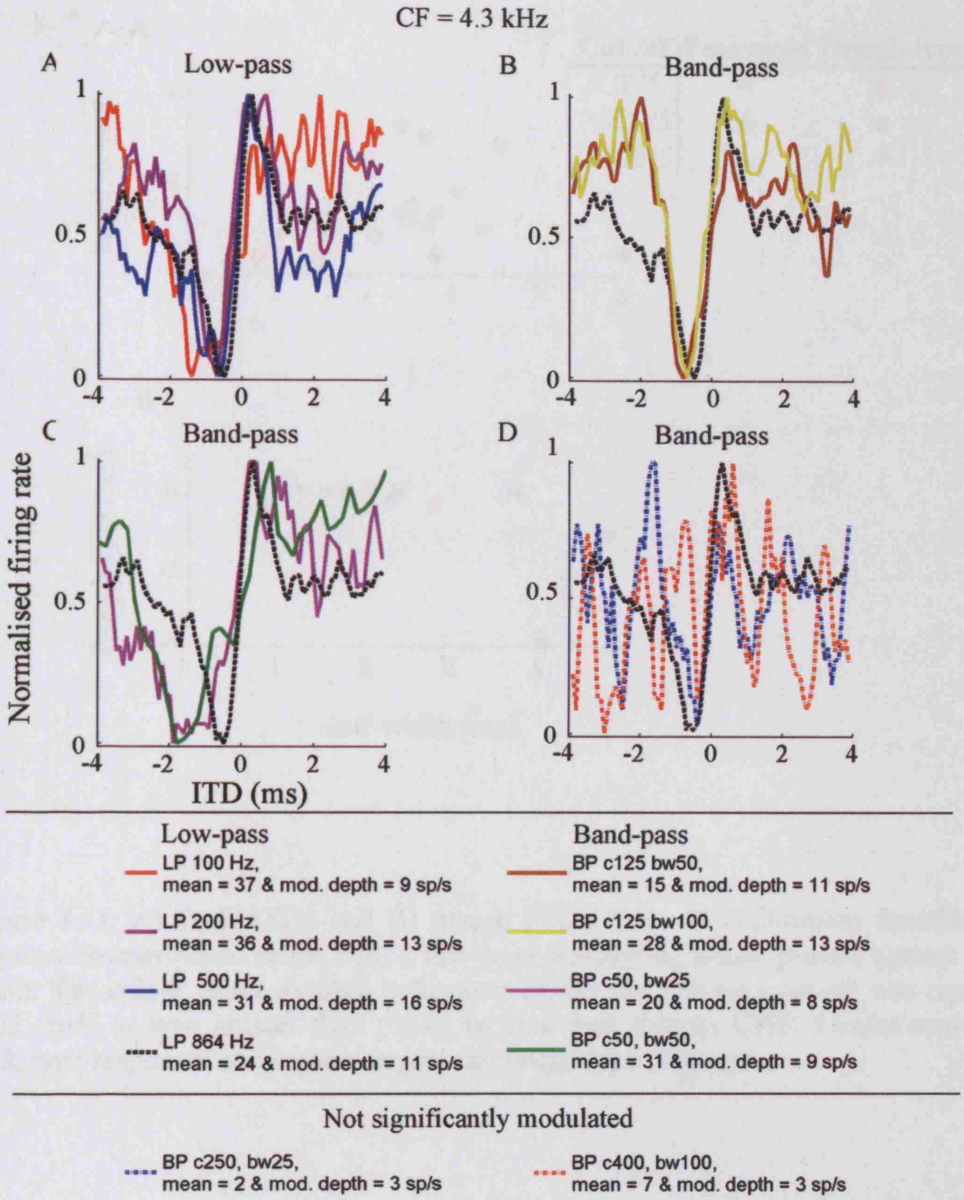

Figure I.14. Normalised ITD-tuning curves generated from the responses of one neuron (neuron 43409, attenuation = 45 dB) in response to transposed noise with the low-pass cut-off of the modulating noise equal to CF/5 (dotted black line) and with different low-pass and band-pass modulating frequencies (coloured lines). The low-pass cut-off (LP) of the stimulus envelope or the centre frequency (c) and the bandwidth of the band-pass (BP) stimulus envelope are indicated in the keys (in Hz). Mean firing rates and modulation depths of the ITD-tuning curves before normalisation are also displayed.

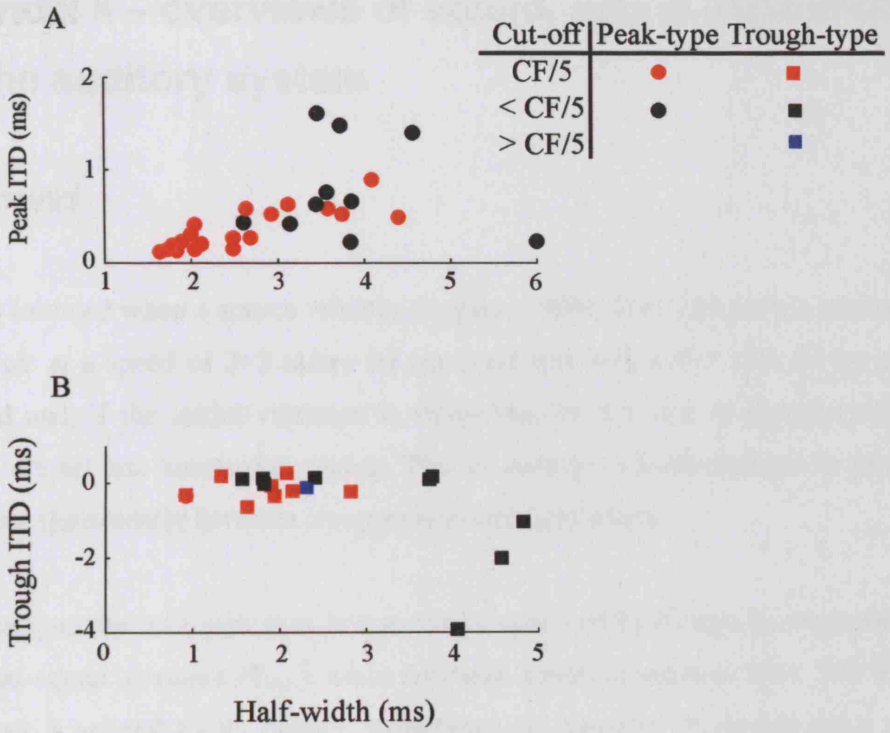

Figure I.15. A) Peak ITDs and B) trough ITDs from all ITD-tuning functions in response to transposed noise with a low-pass modulating noise, plotted against half-width. The colour of the symbol indicates whether the low-pass cut-off was equal to CF/5 (red) or was greater than (blue) or less than (black) CF/5. Circles represent peak-type responses and squares represent trough-type responses.

Appendix II – overviews of sound, sound localisation, and the auditory system

II.1. Sound

Sound is initiated when a source vibrates (Speaks, 1996). This vibration is transmitted through air at a speed of 343 m/sec (at sea level and 20°C). Particles in the air are displaced and, if the initial vibration is sinusoidal (in the case of a pure tone), the particles are set into sinusoidal motion. The air undergoes local changes in pressure, alternating sinusoidally between compression and rarefaction.

The sound pressure of a pure tone is commonly specified by its maximum pressure or root-mean-square pressure (P_{rms}), since its mean sound pressure is zero. The P_{rms} of any sound is related to its power (work/time) or intensity (I; power/area); sound pressure is proportional to the square root of power (and therefore intensity). Equivalently, sound intensity is proportional to the square of the sound pressure. Sound intensity and P_{rms} are normally specified as a decibel level, with reference to a standard sound intensity (I_{ref}) or a standard sound pressure (P_{ref}):

$$dB = 10 \times \log (I/I_{ref}) \quad \text{(Eq. 1)}$$

$$dB = 20 \times \log (P/P_{ref}) \quad \text{(Eq. 2)}$$

A pure tone, a sinusoidally amplitude modulated (SAM) tone and a band-pass, flat spectrum noise are shown diagrammatically in Figure II.1. A pure tone consists of a single sinusoid. Complex sounds can be represented by a Fourier series of sinusoidal components, comprising the addition of single sinusoids (referred to as the sound spectrum). A SAM tone has a spectrum containing three frequencies. The amplitude of the central frequency is 6 dB greater than the other two frequencies, which are of equal amplitude. The central frequency is also referred to as the carrier frequency (f_c). The two "sidebands" are at ($f_c + f_m$) and ($f_c - f_m$), where f_m is the modulation rate. The

phase relationship between the three components is not random; the lowest and highest frequencies are ± θ radians from the central frequency. Noise waveforms are not periodic. A Gaussian noise of infinite duration has a continuous spectrum (in which all frequencies are present) with constant amplitude (also described as a flat spectrum) and random phase. A band-pass flat spectrum noise can be created digitally by setting the amplitude of each spectral component, within the frequency range of the noise, to the same value and by randomly selecting the phase of each spectral component. (The spectral components are determined by the sampling rate and the duration of the noise.)

The envelope of a sound describes changes in amplitude over time and is always positive. The fine-structure of a sound describes the fluctuations in a sound waveform which occur at a faster rate than the sound envelope and cross zero sound pressure. A complex sound waveform can also be represented by a sinusoid with time-varying amplitude and phase (the analytic waveform). The envelope of a narrow-band sound (bandwidth < 0.5 x the centre frequency) is defined by the instantaneous amplitude of this sinusoid (also referred to as the Hilbert envelope; Hartmann, 1998).

The amplitude of a sound waveform is attenuated with increasing distance from a source, due to the impedance of the medium through which the sound is transmitted. When the medium changes, a sound may be reflected, refracted and/or diffracted. When many potentially interfering objects are present, for example the walls of houses, the passage of sound may be considered complex. As will be discussed in Section II.2.2, information about the location of a sound source can be obtained from the way the human body alters a sound waveform.

If two identical sound waveforms arrive simultaneously at each ear, the sound at the ears is referred to as diotic. When the sound waveforms arriving at each ear are different, the sound at the ears is referred to as dichotic. In the simplest case of a sinusoid of frequency *f*, interaural differences can be in intensity or time, and are referred to as interaural intensity differences (IIDs) or interaural time differences

(ITDs), respectively. The ITD of a pure tone can also be described as an interaural phase difference (IPD) in units of degrees, radians or cycles of the period ($1/f$) of the tone.

If a complex waveform is delayed in time (a whole waveform delay), the IPD of each frequency component will be different (see Figure II.2A). In this case, the relationship between IPD and frequency is linear, with a gradient equal to the ITD (or time delay), and an intercept at the origin. The gradient is also called the group delay, and the ITD of each frequency component is called its phase delay. The phase delay of an individual frequency component is equal to the gradient of a line joining its IPD and the origin. When the envelope of a complex waveform is delayed in time but the carrier is not delayed, the envelope-ITD is equal to the group delay. In this case, the linear relationship between IPD and frequency intercepts the abscissa at the carrier frequency (Figure II.2B). Each frequency component has a different phase delay. When the carrier of a complex waveform is delayed in time but the envelope is not delayed, a constant IPD exists across all frequency components, the group delay is zero and the phase delays are different for each frequency component (Figure II.2C). The time delay in the start of a sound is referred to as an onset (or gating) ITD. A sound waveform, when played over headphones, can be constructed to have an onset-ITD which is different from the ongoing-ITD. This is shown diagrammatically in Figure 1.2.

Two time domain waveforms, x and y, sampled uniformly at n points in time, can be compared by calculating the correlation between them. Pearson's correlation coefficient (Eq. 3) is commonly used and is actually the normalised covariance. The normalised correlation (Eq. 4) will be discussed further in Section II.2.9.

Pearson's correlation coefficient =

$$\sum ([x-(\sum x/n)] \cdot [y-(\sum y/n)]) / \sqrt{(\sum [x-(\sum x/n)]^2 \cdot \sum [y-(\sum y/n)]^2)} \qquad \text{(Eq. 3)}$$

$$\text{Normalised correlation} = \sum (x(t) \cdot y(t)) / \sqrt{(\sum x(t)^2 \cdot \sum y(t)^2)} \qquad \text{(Eq. 4)}$$

A cross-correlation function describes the correlation between two waveforms as a function of time delay i.e. the correlation is calculated for a range of time delays imposed between the two waveforms. The correlation between two sinusoids of equal frequency depends on the phase relationships between them. Correlation decreases from 1 to 0 as the IPD increases from 0 to $\pi/2$ radians and continues to decrease to -1 as the IPD increases to π radians. As the IPD increases from π to 2π radians, correlation increases from -1 to 1. A noise can be de-correlated by an IPD (with equal IPD on each frequency component), an ITD, or by adding together correlated and uncorrelated noise in different proportions. A noise that is de-correlated by an ITD is still considered coherent. In auditory psychophysics, coherence is defined as the maximum of a cross-correlation function and can range from 0 to 1. If two otherwise identical signals differ by time, t, their cross-correlation function will be maximal at -t and the signals have a coherence of 1. A noise that is de-correlated by adding together correlated and uncorrelated noise in different proportions has a coherence of less than 1.

The cross-product (Eq. 5) is another metric for quantifying the similarity between two waveforms, as a function of time delay, and differs from the normalised correlation (Eq. 4) because it is not normalised. Cross-products used in Chapters 4 and 7 were calculated can be calculated using Equation 3.

Product (for calculation of cross-products) $= \sum (x(t) \cdot y(t)) / n$ (Eq. 5)

II.2. Sound localisation

The following discussion focuses on human sound localisation. Sound localisation in other mammals will be referred to where appropriate.

II.2.1. Localisation performance

The ability of humans to localise sound can be assessed by asking listeners to indicate the perceived location of a sound, for example using a head-orientating response (Macpherson & Middlebrooks, 2002) or by measuring a threshold separation for discriminating between sound locations, for example by measuring a minimum audible angle [MAA; (Mills, 1958)]. MAAs for pure tones in the horizontal plane, with the reference sound originating from directly in front, can be as low as 1.0° (Mills, 1958). MAAs depend on the nature of the sound presented (Blauert, 1997), the environment in which the measurements are made (Hartmann, 1983) and on the location of the reference sound (Mills, 1958). Sound localisation in the horizontal plane is most accurate when the sound source is directly in front and is less accurate in the vertical plane than in the horizontal plane. Front-back confusions (where a sound is localised to a position which is a reflection of the actual position of the sound about a vertical plane through the ears) are common with narrow-band signals (Blauert, 1997; Middlebrooks & Green, 1991). Sound localisation is most accurate in response to broadband stimuli.

Sound localisation performance has been measured in other mammals including gerbils, cats and monkeys (e.g. Heffner & Heffner, 1988; Heffner & Heffner, 1992; Recanzone & Beckerman, 2004; Heffner & Heffner, 2005; Maier & Klump, 2005; Tollin *et al.*, 2005). It has been suggested that sound localisation performance is related to the width of a mammalian species' field of best vision (Heffner & Heffner, 1992). Data describing the sound localisation performance of guinea pigs have not

been reported [although it appears that guinea-pigs can, with some difficulties, learn various tasks e.g. (Jonson *et al.*, 1975; Evans *et al.*, 1992)].

II.2.2. Cues for sound localisation

Sound transmitted through the air arrives at the eardrums with attributes that depend on the direction of the sound source (Blauert, 1997). Sound emanating from a loudspeaker in free-field (i.e. in an environment where there are no sound reflections) is shown diagrammatically in Figure II.3. In this example, sound arrives at the listener's right ear before the left ear and produces an ITD throughout the duration of the sound, i.e. an ongoing-ITD. The listener forms an obstacle to the progression of a sound wave because of the difference in impedance between the air and the listener. Sound may be diffracted around the head or reflected by the head. Low-frequency sounds are easily diffracted whereas higher frequency sounds are reflected away from the head, creating an IID with greater intensity, in this example, at the right ear. The shape of the external ear (composed of the pinna and the ear canal) also influences the sound at the eardrum. The external ear shows resonances that depend on the location of the sound source; some sound frequencies are boosted or attenuated relative to other frequencies. This filtering provides spectral cues for localisation of a sound source. Finally, sound reflections from the rest of the body will also influence the sound arriving at the eardrum. In general, the cues for sound localisation can be divided into three categories: 1) ITDs, 2) IIDs and 3) spectral cues. Cues relating to differences in the spectrum of a sound are often referred to as "monaural" cues. This is because spectral cues may provide localisation information even if the listener is deaf in one ear and not because they are only used monaurally (Martin *et al.*, 2004).

The ITDs, IIDs and spectral cues for sound localisation can be measured close to the eardrum by placing small probe microphones inside the ear canals. The head-related transfer function (HRTF) relates the sound presented from a loudspeaker to the sound measured at the eardrum and is calculated as a function of the location of the

loudspeaker. Interaural transfer functions describing differences in sound between the two ears are derived from HRTF measurements. Both IIDs and ITDs increase with the horizontal displacement of a sound and are maximal at ~90° from the midline (Blauert, 1997). Maximum IIDs occur in sounds above 5 kHz and can be as large as 35-40 dB (Shaw, 1974; Middlebrooks *et al.*, 1989). Above ~8 kHz, IIDs also vary with sound elevation due to the resonances of the pinnae. ITDs are less dependent on sound frequency but are largest at frequencies below 1 kHz (Kuhn, 1977). ITDs are maximally ~1 ms (Blauert, 1997). ITDs in the envelope of HRTF measurements at frequencies between 4 and 12 kHz can provide directionally-dependent ITD cues up to a maximum ITD of ~1 ms (Middlebrooks & Green, 1990). Spectral cues for sound localisation predominantly occur at frequencies above 2 kHz (Shaw, 1974).

According to the duplex theory of sound localisation, ITDs are a cue for localising low-frequency sounds (< 1.5 kHz) and interaural intensity differences (IIDs) are a cue for localising high-frequency sounds (Stevens & Newman, 1936). Stevens & Newman (1936) measured sound localisation in free-field in response to pure tones and observed that performances using tones between 2 and 4 kHz were particularly poor. This was explained by weak cues for localisation at these frequencies. It is not possible to detect ITDs in pure tones with frequencies > 2 kHz, and IIDs are relatively small for sounds between 2 and 4 kHz, explaining the poor performance in this region. Mills showed that MAAs in response to low-frequency tones correspond closely to thresholds for detection of ITDs and that MAAs at high-frequencies correspond closely to thresholds for detection of IIDs, supporting the duplex theory (Mills, 1960). The duplex theory has been shown to hold for complex sounds (Macpherson & Middlebrooks, 2002). Throughout this thesis "high-frequency" refers to sounds ≥ 2 kHz, i.e. frequencies at which sensitivity to ITDs in pure tones does not occur. Low-frequency, in this thesis, refers to frequencies of pure tone at which sensitivity to ITDs occurs.

The directionally-dependent ITDs, IIDs and spectral cues that are available to a listener depend on the size and shape of the listener's head and pinnae. The maximum

possible ITDs and IIDs will be large for an animal with a large head size (Musicant *et al.*, 1990). The frequency range of hearing will also affect which cues are used for localisation. Guinea pigs have a hearing range comparable to that of humans (Fay, 1988). Sterbing *et al.* (2003) recorded responses to broadband noise using microphones placed inside the ear canals of 18 guinea pigs and found the maximum IID to be 15-20 dB and the maximum ITD to be 330 μs at around 400-600 Hz. A theoretical consideration of the time taken for sound to travel around a spherical head gives a maximum ITD of ~150 μs for the guinea pig (Kuhn, 1977; McAlpine *et al.*, 2001). Sterbing *et al.* (2003) account for the large difference in these values by assuming that the diffraction pattern of sound around the head and pinnae creates a longer travel time than that predicted by a spherical head model. They did not calculate the size of envelope ITDs within high-frequency sounds.

It is also noted that head (or pinnae) movements can be used to localise sounds and that visual cues to the localisation of a sound source can influence sound localisation (Blauert, 1997). Human sensitivity to IIDs, ITD and spectral cues is discussed in the following three Sections. ITDs, which are the topic of this thesis, will be discussed in depth.

II.2.3. Sensitivity to spectral cues

If a sound is recorded by one microphone and played back to a listener diotically over headphones, the resulting sound image is heard as though it originates from within the head. When a stereophonic recording is presented over headphones, the resulting sound image remains within the head and its lateral position within the head will depend on the ITDs and IIDs present. Studies presenting sound over headphones, and in which the sound image remains within the head (such as the study described in Chapter 5), are referred to as studies of *lateralisation*. This is in contrast to studies measuring the *localisation* ability of sounds originating in the free-field. If sound is recorded at each eardrum and is played back to the listener over stereophonic head-

phones the sound is externalised (Plenge, 1974; Wightman & Kistler, 1989a; Wightman & Kistler, 1989b; Hartmann & Wittenberg, 1996). Sounds can be externalised when the spectral distortions created by the pinnae are only crudely maintained (Kulkarni & Colburn, 1998). Sounds that are presented via headphones, and that have been filtered according to a subject's HRTF in order to simulate sound from a source in the free-field, are localised in what is referred to as Virtual Acoustic Space (VAS). The inability of listeners to discriminate between free-field sound and VAS has been used as verification of the correct measurement and reproduction of free-field sound e.g. Wightman & Kistler (1989b).

As well as producing externalisation of a sound image, spectral cues help to localise sounds with ambiguous ITDs or IIDs. Regions of constant ITD or IID describe a "cone of confusion" around the head (Blauert, 1997). For example, sounds on the median plane (at the midline) have 0-μs ITD and 0-dB IID, irrespective of the elevation of the sound source (although see Blauert, 1997, page 95). The pinnae are important for localisation in the median plane e.g. as shown by blocking the hollows in the pinnae but leaving the entrance to the ear canals open (Gardner & Gardner, 1973). A listener's individual spectral cues can be used to resolve front-back confusions in the location of a sound (Wenzel *et al.*, 1993). As might be expected given that spectral cues consist of relative boosts and attenuations in sound frequencies, they are most useful for the localisation of a broadband sound. Studies of vertical localisation at the midline, where ITDs and IIDs are approximately constant (although see Blauert, 1997) can be used to assess sensitivity to spectral cues only. MAAs in the median plane can be as low as 4° (Blauert, 1997).

The use of VAS allows separate manipulation of spectral cues, ITDs and IIDs. Wightman & Kistler (1997) presented VAS stimuli monaurally and concluded that monaural performance using spectral cues was very poor. This is not to conclude, however, that monaural spectral cues are not important under binaural listening

conditions for localisation in the vertical plane and for resolving front-back confusions (Martin *et al.*, 2004).

II.2.4. Sensitivity to Interaural Intensity Differences

Thresholds for detection of an IID, from a reference of 0-dB IID, have been measured using a variety of signals (Blauert, 1997). Thresholds are obtained in headphone studies by measurement of the smallest IID that can be detected. The sounds are not externalised and, apart from the IID, are diotic. As the IID is increased, the sound image moves towards the ear at which it is louder. Thresholds for detection of an IID are around 1 dB and are relatively constant over a broad range of frequencies (Mills, 1960), in contrast to sensitivity to ITDs (see Section II.2.5). Sensitivity to IIDs is decreased around 1 kHz (Grantham, 1984). As sound frequency decreases the IID present in a sound emanating in free-field decreases and so the IID threshold corresponds to an increasingly large angular distance from the midline. IIDs therefore become less informative as frequency decreases. Sounds are fully lateralised (to a position close to one ear) at around 15-20 dB IID (Blauert, 1997).

II.2.5. Sensitivity to Interaural Timing Differences

ITD thresholds for pure tones (from a reference of 0-μs ITD) have been measured as a function of sound frequency. It becomes impossible to detect ITDs in pure tones above ~1.5 kHz (Yost, 1974; Klumpp & Eady, 1956; Zwislocki & Feldman, 1956). Thresholds are minimal for pure tone frequencies of ~1 kHz (Klumpp & Eady, 1956; Zwislocki & Feldman, 1956). Thresholds for discrimination of changes in ITD have been measured as a function of the reference ITD. Thresholds are smallest from a reference of 0-μs ITD and largest from a reference towards either ear [at an IPD of 180° (Yost, 1974; Hafter *et al.*, 1975; Domnitz & Colburn, 1977)]. There is no equivalent of the visual fovea (the small portion of the retina over which spatial acuity is greatest and over which attention is usually directed). Discrimination of

ITDs has also been measured as a function of stimulus duration and for durations up to ~300 ms thresholds improve, beyond which they remain constant (Tobias & Zerlin, 1959; Hafter *et al.*, 1979; Houtgast & Plomp, 1968; Buell & Hafter, 1988). As a function of intensity, thresholds for detection of ITDs in a pure tone decrease rapidly from threshold and remain relatively constant beyond 20-30 dB sensation level (Zwislocki & Feldman, 1956). There is a reduction in sensitivity (an increase in threshold) at sensation levels above ~70 dB, particularly for pure tones $\geq$ 1 kHz (Zwislocki & Feldman, 1956).

In the late 1950s, several authors described thresholds for detection of ITDs in high-frequency, complex sounds [SAM tones: (Leakey *et al.*, 1958), high-pass filtered clicks: (David *et al.*, 1959), high-frequency noise bands: (Klumpp & Eady, 1956)]. Further investigations using SAM tones (Henning, 1974a; Henning & Ashton, 1981; Henning, 1980; Nuetzel & Hafter, 1981), bands of noise (McFadden & Pasanen, 1976) and two-tone complexes (McFadden & Pasanen, 1976) demonstrated that sensitivity to ITDs within high-frequency sounds is conveyed by the extraction of ITDs in the envelope of high-frequency sounds. Sensitivity to ITDs within a high-frequency stimulus required the presence of relatively low-frequency amplitude fluctuations.

Sensitivity to changes in ITD conveyed by high-frequency stimuli has typically been found to be poorer (i.e. thresholds are higher) than that measured with low-frequency stimuli (Yost *et al.*, 1971; Bernstein & Trahiotis, 1982; Jones & Williams, 1981). The lateral extent of the intracranial image produced by an ITD conveyed by high-frequency complex stimuli is typically smaller than that conveyed by low-frequency stimuli (Bernstein & Trahiotis, 1985b; Trahiotis & Bernstein, 1986). There is also an upper cut off in the f_m at which sensitivity to ITDs can occur (McFadden & Pasanen, 1976; Nuetzel & Hafter, 1981; Bernstein & Trahiotis, 1994). Bernstein & Trahiotis (1994) note that envelope ITD sensitivity decreased markedly at a carrier frequency of 12 kHz. Sensitivity to ITDs in the envelope of high-frequency sounds is greatest for f_m around 125 Hz. (This is a low f_m compared to the 1 kHz pure tone frequency at

which discrimination of ITDs in pure tones is most sensitive.) Sensitivity to ITDs in high-frequency sounds is also more dependent on intensity than in low-frequency sounds, and is better at higher intensities (McFadden & Pasanen, 1976; Nuetzel & Hafter, 1976; Simon *et al.*, 1994).

More recently, Bernstein and Trahiotis (2002) have shown, using transposed tones, that sensitivity to ITDs within the envelope of high-frequency waveforms can be equivalent to that seen at low-frequencies. This occurs at relatively low f_m, around 126 Hz and there is still an upper limit on the f_m at which sensitivity to ITDs occurs. Transposed tones were designed to provide auditory nerve fibres (ANFs) that respond to high-frequency sounds with a probability of action potential firing similar to that observed in response to low-frequency tones (see Section 1.1). Bernstein & Trahiotis (2003) demonstrated that extents of lateralisation achieved with high-frequency transposed noise stimuli could be equivalent to lateralisation observed with low-frequency noise stimuli.

II.2.6. Combined cues

Headphone studies examining sensitivity to ITDs and IIDs provide insight into the relative importance of these cues for sound localisation. However, studies using ITDs and IIDs alone are inherently limited; the percept that is under investigation is of a sound image that remains inside the head (lateralised and not localised) and the cue under investigation is examined when it is in conflict with all other cues (which are held constant). Variability in lateralisation has been shown to be minimal when ITD and IID cues are most "natural" (Gaik, 1993), i.e. when ITDs and IIDs correspond to the same location in space.

Investigators have tried to compare the salience of ITDs and IIDs for lateralisation by measuring time-intensity trading ratios (David *et al.*, 1959; Harris, 1960). The ITD necessary to centre a sound image at the midline, of a sound with a set IID, is

measured for a range of IIDs. The gradient of the resulting function provides the time-intensity trading ratio (in μs per dB), describing the ITD required to centre a stimulus with an IID that otherwise mediates a lateral sound image. Time-intensity trading ratios are generally higher at high-frequencies than at low-frequencies (Harris, 1960), suggesting that ITDs are less salient within high-frequency sounds. However, these experiments assume that ITDs and IIDs can be traded whereas there is evidence to the contrary (Hafter & Carrier, 1972). In experiments measuring time-intensity trading ratios, the perception of two sound images has also been described, one which has a greater time-intensity trading ratio than the other (Whitworth & Jeffress, 1961; Hafter & Jeffress, 1968).

The relative importance of ITDs and IIDs has also been compared using VAS. ITDs at low-frequencies have been demonstrated to dominate human sound localisation (Wightman & Kistler, 1992) and ITDs at high-frequencies have been shown to be relatively unimportant for sound localisation (Macpherson & Middlebrooks, 2002).

II.2.7. Localisation in reverberant environments and onset-ITDs

It is a common experience that, when listening in a reverberant environment, sounds are correctly localised towards the sound source. This occurs despite the arrival at the eardrums of reflections which may provide different spatial cues [although localisation performance does degrade slightly in reverberant environments (Hartmann, 1983)]. The dominance for localisation of the first arriving waveform is referred to as the precedence effect (Wallach *et al.*, 1949; Gardner, 1968; Litovsky *et al.*, 1999). In highly reverberant environments, the presence of many interacting reflections may be considered to generate an uncorrelated background noise over which the signal from the direct source must be detected.

The precedence effect is strongest for sound having a fast onset and therefore a spread of spectral components (Hartmann, 1983). Following the example set by

Wallach *et al.* (1949), studies relating to the precedence effect have often been conducted using pairs of clicks (Yost & Soderquist, 1984; Saberi & Perrott, 1990; Freyman *et al.*, 1997; Litovsky *et al.*, 1999). The notion that directional information at the onset of a stimulus (e.g. in the first click of a click pair) dominates localisation has led to the hypothesis that subsequent directional information (e.g. in the second click pair) is suppressed by neural inhibition (Zurek, 1980; Houtgast & Aoki, 1994). However, data which suggest a build-up and break-down of the precedence may be explained by high-order processes (Freyman *et al.*, 1991; Litovsky *et al.*, 1999). Other authors emphasise the essentially peripheral nature of onset dominance and suggest that the dominance of the first arriving click pair can be explained by interactions between clicks on the basilar membrane (Hartung & Trahiotis, 2001). Various paradigms have been used to study the precedence effect (Litovsky *et al.*, 1999) and it is likely that more than one mechanism will account for the data obtained.

The salience of onset-ITDs in the discrimination of ITDs within low-frequency tones and noise has been investigated (Tobias & Schubert, 1959; Buell *et al.*, 1991). These authors concluded that an onset-ITD has little influence on discrimination of ongoing-ITDs. Buell *et al.* (1991) suggested that onset-ITDs, at low-frequencies, are used when the ongoing-ITD information is ambiguous. They showed that an onset-ITD can enable listeners to disambiguate an ongoing-ITD equal to 0.5 cycles of a low-frequency tone stimulus. At 0.5 cycles, an ITD can be localised towards either the left or right ear with equal probability. The percept can switch spontaneously between the two positions. An onset-ITD favouring the left ear can stabilise the percept towards that ear (Buell *et al.*, 1991).

Few studies have measured the salience of onset-ITDs in high-frequency sounds. MacPherson & Middlebrooks (2002) suggest that onset-ITDs in high-frequency sounds contribute little towards sound localisation. Thresholds for discrimination of onset-ITDs have been measured using high-frequency clicks and have been found to be < 100 μs for experienced listeners (Buell & Hafter, 1988). Using high-frequency

click trains, Hafter & Buell (1990) show that, for click rates greater than 80 s^{-1}, more information about the ITD is gained from the first click than from subsequent clicks. However, this does not necessarily suggest that the onset-ITD dominates lateralisation. Zurek (1980) hypothesised that onset-ITDs are more salient for discrimination of ITDs within low- as compared to high-frequency sounds. Experiments investigating the salience of onset-ITDs for discrimination of ITDs in SAM tones are described in Chapter 5.

II.2.8. Binaural advantage in signal detection

In a highly reverberant environment, or in the presence of many sounds from different sources, the auditory system faces the problem of detecting a salient signal within background noise. An example of this problem is the "cocktail party effect" where it is possible to attend to the voice of one person to the exclusion of competing sounds. This can be thought of as signal detection in noise, as sound source segregation (i.e. how are different sources of sound separated) or as auditory grouping (i.e. what components of the sound at the ears are from the same sound source).

Psychophysically, the role of ITDs in improving signal detection has been investigated by measuring binaural masking level differences (BMLDs). BMLDs are usually measured for the detection of a tone masked by a broad-band noise [as first studied by (Hirsh, 1948)]. A BMLD is the difference between the signal level at threshold of detection of a signal and the signal level in a reference condition (Webster, 1951). The reference condition is usually the masker level at threshold for detection of a signal when the masker and signal are presented to the same ear (Blauert, 1997). In the diotic condition the BMLD is around 0 dB, indicating that there is little advantage in detection of the signal above the monaural case. However, if a 500-Hz tone is anti-phasic (i.e. with an IPD of π) BMLDs in broadband noise can be as large as 15 dB. This condition is referred to as the N_0S_π condition (noise diotic,

signal in π phase). If the 500-Hz tone is diotic and the broadband noise masker is presented in π phase (the $N_\pi S_0$ condition), BMLDs can be up to 12 dB. Similar to the BMLD, improvements in the intelligibility of speech are observed when speech masked by diotic noise is presented with an ITD other than zero μs [the binaural intelligibility level difference; (Licklider, 1948)].

BMLDs have been recorded in a variety of conditions (Jeffress *et al.*, 1952; Jeffress *et al.*, 1962; Robinson & Jeffress, 1963; Langford & Jeffress, 1964) and attempts have been made to directly relate BMLDs to the difference in ITDs of the signal and masker (Licklider, 1948; Hafter *et al.*, 1969). Sounds with the same ITD are more likely to originate from the same source (although note that the presence of reflections complicates this situation) and therefore ITDs might be useful in grouping auditory objects and in improving the detection of an auditory object in the presence of background noise. However, this model does not explain all BMLD data. For example, BMLDs can be measured for high-frequency pure tones which cannot be localised on the basis of ongoing-ITDs. BMLDs at measured using high-frequency sounds have been explained by the detection of changes in envelope correlation caused by the presence of the tone (Bernstein & Trahiotis, 1992).

A number of authors have specifically investigated BMLDs within the envelope of high-frequency sounds and have reported lower BMLDs than those observed at low frequencies (Henning, 1974b; McFadden & Pasanen, 1978; Bernstein & Trahiotis, 1992). However, van der Par & Kolhrausch (1997) showed that sensitivity to BMLDs at high frequencies can be equivalent to those at low frequencies by using "transposed" stimuli.

II.2.9. Models based on cross-correlation

Models based on cross-correlation have been successful in accounting for sensitivity to ITDs, lateralisation and for signal detection in noise (Stern & Trahiotis, 1997).

Such models assume that the sound at each ear is cross-correlated by the auditory system and consist of a τ-f plane which can be considered to represent neurons that respond maximally to different sounds frequencies and different ITDs. This can be considered a mathematical implementation of Jeffress' 1948 model (see Section 1.3). The firing rates of a τ-f plane of neurons have been modelled, at each frequency, by the cross-correlation of the sound at each ear, following peripheral processing (Colburn, 1973; Stern & Trahiotis, 1997). Trahiotis & Stern (1989) and Stern *et al.* (1988) extended the τ-f matrix of neurons to predict the lateralisation of a range of sounds. Although models based on cross-correlation appear to be powerful in their explanation of psychophysical data, it is noted that evidence for the existence of a τ-f plane in the mammalian central auditory system is, at best, equivocal (see Section 1.3).

van de Par & Kohlrausch (1995) suggest that the measure of correlation used will affect the results of cross-correlation models. For waveforms that have a mean of zero, Pearson's correlation (Eq. 3, Section II.1) is equal to the normalised correlation (Eq. 4, Section II.1). For waveforms that do not have a mean of zero, Pearson's correlation and the normalised correlation are different. Modelling using the normalised correlation can account for a wide range of psychophysical data on sensitivity to ITDs, BMLDs and lateralisation (Bernstein & Trahiotis, 1996a; Bernstein & Trahiotis, 1996b). Bernstein and Trahiotis (2002) calculate the normalised correlation as a function of ITD of transposed tones, after peripheral processing, to account for human psychophysical data. The threshold for detection of an ITD, at a given f_c, is assumed to occur at a constant value of normalised correlation.

II.3. The auditory system

A brief overview of the mammalian auditory system is outlined in Section II.3.1. In Sections II.3.2-12, different parts of the auditory system are reviewed in more detail. The focus of each review will be on anatomy and physiology with respect to detection of cues for sound localisation and particularly ITDs. Much of the information in these Sections was obtained from textbooks (Popper & Fay, 1992; Webster *et al.*, 1992).

II.3.1. An overview

A diagram of the mammalian auditory system is shown in Figure II.4. Sound is transduced into electrical activity in the cochlea. Action potentials in auditory nerve fibres (ANFs) excite neurons in the cochlear nucleus which contains a variety of cell types. From the cochlear nucleus, parallel ascending pathways transform the input originating from the auditory nerve. The first site of binaural interaction, where neurons integrate information from both ears, occurs in the superior olivary complex (SOC); although there are responses to contralateral sound in the cochlear nucleus (CN; Shore *et al.*, 2003). The textbook view of the SOC is that ITDs are processed in the medial superior olive (MSO) and IIDs are processed in the lateral superior olive (LSO). The nuclei of the lateral lemniscus receive inputs from both monaural and binaural pathways and send projections, along with those from the SOC and CN, to the inferior colliculus (IC). Forming synapses in the IC has been considered obligatory for ascending projections of the auditory system. Neurons in the IC project to the medial geniculate nucleus (MGN) which forms the auditory thalamus. Neurons in the MGN, in turn, project to the primary auditory cortex. Descending auditory pathways, including the olivo-cochlea pathways, exist but these will not be discussed.

II.3.2. The cochlea and sound transduction in the inner ear

An understanding of the transduction of sound into activity in auditory nerve fibres (ANFs) is essential to understanding the auditory system and the design of the transposed sound stimuli used in this thesis. The anatomy and function of the cochlea are described below.

Sound waves vibrate the tympanic membrane (ear drum), setting in motion the three bones of the middle ear (incus, malleus and stapes). The stapes is coupled to the oval window of the cochlea and transmits the vibration to the fluid (perilymph) in the scala vestibuli inside the cochlea. The scala vestibuli is continuous with the scala tympani via an opening (the helicotrema) at the apical end of the cochlea. Throughout most of the cochlea, the scala vestibuli and scala tympani are separated by the basilar membrane and Reissner's membrane. Perilymph is sealed inside the scala tympani by the flexible round window. The basilar membrane is deflected towards and away from the scala vestibuli as a result of pressure differences in the scala vestibuli and scala tympani caused by vibration at the oval window. Von Bekesy, working with cochleae from human cadavers, first described a travelling wave of vibration along the basilar membrane in response to sinusoidal vibration of the stapes (von Bekesy, 1947). The maximal deflection occurs at a specific point on the basilar membrane where its resonance matches the frequency of the stimulating tone. The basilar membrane vibrates maximally to high-frequency sounds at its base, which is stiffer and less elastic than at its apex (Robles & Ruggero, 2001).

The organ of Corti sits on the basilar membrane and moves with it. Inside the organ or Corti are hair cells, so named because of bundles of stereocilia on their apical surface. The stereocilia on each hair cell are arranged in rows that are graded in height (Hudspeth, 2005). Movement of the basilar membrane towards the scala vestibuli deflects stereocilia towards the tallest row. This places tension on "tip links" connecting the stereocilia together, causing cation-selective ion channels in the

membrane of the stereocilia to open. The resulting influx of K^+ and Ca^{2+} ions leads to depolarisation of the hair cell. Movement of the basilar membrane towards the scala tympani deflects stereocilia towards the shortest row and releases tension in the tip links, closing the ion channels which they gate; the hair cell is hyperpolarised. The membrane potential of a hair cell reflects the vibration of the basilar membrane caused by a sound waveform (Fettiplace & Fuchs, 1999). The magnitude of hair cell depolarisation depends upon the magnitude of the basilar membrane motion and hair cells are therefore tuned to sound frequency according to their position along the basilar membrane.

Hair cells have sharper tuning for sound frequency than can be accounted for by the "tuning" of the basilar membrane in metabolically-inactive cochleae (from dead animals). Outer hair cells (OHCs) change their length in response to sound (Ashmore, 1987) and it is believed that active lengthening and shortening of OHCs amplifies the vibration of the basilar membrane, sharpening tuning to sound frequency and leading to the production of otoacoustic emissions Kemp (1978). The stereocilia themselves may also contribute to the active amplification in the cochlea (Kennedy *et al.*, 2005). Active processes generate a compressive non-linearity in the movement of the basilar membrane (Rhode, 1971; Sellick *et al.*, 1982; Ruggero *et al.*, 1997; Recio *et al.*, 1998).

A second class of hair cells, the inner hair cells (IHCs), generate action potentials in type I ANFs, which are believed to be the primary afferents of the cochlea (see section II.3.3). In response to low-frequency tones, the membrane potential of an IHC oscillates with the period of the tone. The membrane potential is asymmetric in that it is more depolarised in response to a low-frequency tone than hyperpolarised. The upper rate of fluctuation of IHC membrane potentials is limited by their membrane time constant (equal to capacitance x resistance). As tone frequency increases the oscillation of guinea pig IHC membrane potentials decrease in magnitude in relation to a depolarised, dc component to the potential. In response to tones > ~1 kHz, the

membrane potential of guinea pig IHCs are dominated by a depolarised, dc component (Palmer & Russell, 1986).

II.3.3. Auditory nerve fibres (ANF)

Anatomy

Two types of ANFs have been identified. Type I ANFs are myelinated and synapse with a single IHC (Ryugo, 1992). These myelinated fibres transmit action potentials rapidly towards the cochlea nucleus. Type II ANFs synapse with >1 OHCs (Ryugo, 1992). These unmyelinated fibres are not thought to be part of the primary ascending auditory system and therefore only Type I ANFs will be discussed.

Physiology

In response to a continuous stimulus, action potential discharge in ANFs adapts from an initially high rate to a lower rate (Yates *et al.*, 1985). ANFs can be characterised by their spontaneous firing rate, by the frequencies of sound to which they respond and by the lowest sound intensity (threshold) at which action potentials can be evoked. The sound frequency for which threshold is lowest defines a neuron's characteristic frequency (CF). The frequency-tuning of ANFs occurs as a result of the position along the basilar membrane of the IHC which they innervate (Ruggero, 1992). Latencies from sound onset to action potential firing are shorter in response to high-frequency than to low-frequency stimuli because of the travel time of the basilar membrane travelling wave and are also shorter with increasing intensity (Heil & Neubauer, 2003). The filter characteristics of an ANF can be expressed as an equivalent rectangular bandwidth (ERB; Moore, 2004). The ERB of ANFs increases with increasing frequency (Evans *et al.*, 1992). Above ~4 kHz, guinea pig ERBs are narrower than those of humans (measured behaviourally) and below ~4 kHz, guinea pig ERBs are broader than those of humans (Evans *et al.*, 1992).

The probability of action potential firing in an ANF increases when the IHC it synapses with is depolarised. IHC membrane potentials oscillate in response to low-frequency tones and, therefore, action potentials in ANFs are locked to the period of a low-frequency tone (Tasaki, 1954; Rose *et al.*, 1967). In all mammalian species studied so far, there is a cut-off frequency at which phase-locking in ANFs deteriorates (Weiss & Rose, 1988). This reflects the reduction in the oscillatory component of IHC membrane potentials relative to the dc component (see section II.3.2; Palmer & Russell, 1986). In guinea-pig ANFs, phase-locking deteriorates at ~1 kHz (Palmer & Russell, 1986) whereas in cats the cut-off is higher at ~2 kHz (Johnson, 1980).

ANFs have also been recorded in response to complex sounds. In response to low-frequency stimuli, the firing rate of an ANF is stimulus-locked and can be approximated by the half-wave rectified stimulus (Brugge *et al.*, 1969). In response to high-frequency complex stimuli, firing rate of an ANF increases with increasing intensity of the stimulus and follows the envelope of the stimulus. Action potentials are therefore phase-locked to the sinusoidal envelope of high-frequency SAM tones (Palmer, 1982; Joris & Yin, 1992). Phase-locking to SAM tones has a cut-off frequency at which phase-locking decreases. Palmer (1982) suggests that this is due to cochlear filtering of the spectral side-bands of a SAM tone. Joris and Yin (1992) suggest that there may be an additional limitation on the modulation frequencies (f_m) at which phase-locking to high-frequency, amplitude modulated sounds can occur in ANFs. Responses to amplitude modulated stimuli are compressed compared to a sound stimulus due to the active processes in the cochlea which create non-linear movement of the basilar membrane as a function of sound intensity (Joris & Yin, 1992). ANF responses to broadband noise are stimulus-locked and reflect their band-pass tuning for sound frequency (Louage *et al.*, 2004). In response to a click, ANFs respond for longer than the duration of the click, particularly at low-frequencies, where the frequency tuning of ANF is relatively narrow (e.g. Versnel *et al.* 1992).

Sensitivity to ITDs requires sensitivity to the temporal structure of a sound waveform. Sensitivity to ITDs can occur in response to the carrier of low-frequency sounds because phase-locking can occur in response to low-frequency sounds. Sensitivity to ITDs occurs in response to the amplitude modulated envelope of high-frequency sounds because phase-locking can occur in response to the envelope of high-frequency sounds.

II.3.4. The cochlear nucleus (CN)

Anatomy

ANFs enter the cochlear nucleus and bifurcate, sending one branch to the anterior ventral cochlear nucleus (aVCN) and the other branch to the posterior ventral cochlear nucleus (pVCN) and the dorsal cochlear nucleus (DCN) (Ryugo, 1992). The projections of ANFs are tonotopically organised (Ruggero, 1992). The pVCN contains octopus cells which respond to broadband input and are specialised for the detection of well-timed synchronous activity (Golding *et al.*, 1995; Oertel *et al.*, 2000). The DCN has a layered architecture and neural circuits within the DCN are thought to generate sensitivity to spectral cues for sound localisation (Young & Davies, 2001).

In the aVCN, ANFs synapse with bushy cells which in turn project to neurons in the SOC that are sensitive to ITDs and IIDs (Yin, 2001). Bushy cells have compact dendrites and somata which are either spherical (spherical bushy cells; SBCs) or oblong (globular bushy cells; GBCs; Cant, 1992). A few (1-3) ANFs synapse onto each SBCs forming giant synapses called endbulbs of Held, whereas GBCs receive a higher number of ANF inputs via modified endbulbs (Cant, 1992).

SBCs project to the ipsilateral LSO and to the ipsilateral and contralateral MSO (Smith *et al.*, 1993). GBCs project to the contralateral medial nucleus of the trapezoid body (MNTB) where they form the calyx of Held (Smith *et al.*, 1991).

Physiology

Post stimulus time histograms (PSTHs) from SBCs and GBCs are described as primary-like and primary-like with notch, respectively, to reflect their similarity to firing in ANFs (Rhode & Greenberg, 1992). The responses of bushy cells are locked to a phase of the period of a low-frequency tone and phase-locking is enhanced in comparison with responses in ANFs (Joris *et al.*, 1994a; Joris *et al.*, 1994b; Louage *et al.*, 2005). The enhanced phase-locking is likely to be the result of coincidence detection of more than one ANF input (Rothman *et al.*, 1993; Joris *et al.*, 1994a). Bushy cells have membrane properties, including specific K^+ channels that enable them to follow rapid firing in ANFs (Rothman *et al.*, 1993; Rothman & Manis, 2003).

II.3.5. Superior Olivary Complex (SOC)

The SOC is considered to be the first site in the ascending auditory system where binaural interaction occurs and has long been implicated in the detection of ITDs and IIDs for sound localisation (Goldberg & Brown, 1969). When the SOC is lesioned, sound localisation performance is disrupted in both left and right hemifields (Kavanagh & Kelly, 1992). Only the MNTB, medial superior olive (MSO) and lateral superior olive (LSO) will be discussed below although it is noted that the SOC does contain other nuclei, including the ventral nucleus of the trapezoid body (VNTB).

II.3.6. Medial Nucleus of the Trapezoid Body (MNTB)

Anatomy

The MNTB contains principal neurons which are contacted by the largest synapse in the brain, the calyx of Held (Forsythe, 1994). The calyx of Held is formed from the termination of GBC axons from the aVCN and is excitatory (Forsythe & Barnes-Davies, 1993). The MNTB provides inhibitory glycinergic inputs to the LSO and to the MSO (Smith *et al.*, 1998).

Physiology

Action potentials of MNTB principal neurons are phase-locked in response to pure tones (Smith *et al.*, 1998). It has been suggested that phase-locking becomes more precise between the GBC inputs and the firing of MNTB neurons (Kopp-Scheinpflug *et al.*, 2003). Specialised K+ conductances enable MNTB neurons to follow fast trains of action potentials in their SBC inputs (Brew & Forsythe, 1995; Dodson *et al.*, 2002).

II.3.7. Medial Superior Olive (MSO)

Anatomy

MSO principal neurons are bipolar with medial dendrites that receive excitatory inputs from contralateral SBCs and lateral dendrites that receive excitatory inputs from ipsilateral SBCs (Smith *et al.*, 1993; Smith, 1995; Cant & Benson, 2003). MSO neurons also receive inhibitory inputs from neurons in the MNTB and VNTB (Grothe & Sanes, 1993; Smith, 1995). In the gerbil, inhibitory inputs undergo a developmental shift from covering the entire surface of the neuron to being localised specifically to the soma (Kapfer *et al.*, 2002; Seidl & Grothe, 2005).

MSO principal neurons are arranged tonotopically within the MSO and, in mammals that use ITDs to localise sounds, there is an over-representation of neurons that respond to low-frequency sounds (Goldberg & Brown, 1969; Guinan *et al.*, 1972). The MSO sends excitatory projections ipsilaterally to the inferior colliculus and the nuclei of the lateral lemniscus (Oliver, 2000).

Physiology

It is reportedly difficult to record responses of single MSO neurons *in vivo* (Goldberg & Brown, 1969; Guinan *et al.*, 1972; Caird & Klinke, 1983). This is likely due to the large neurophonic potential recorded within the MSO and the small amplitude of action potentials at their soma (Scott *et al.*, 2005). Studies that have successfully recorded from single neurons in the MSO demonstrate that the firing rates of MSO neurons are modulated by ITDs [cat (Galambos *et al.*, 1959; Caird & Klinke, 1983; Yin & Chan, 1990), dog, (Goldberg & Brown, 1969), kangaroo rat (Moushegian *et al.*, 1975; Crow *et al.*, 1978), gerbil: (Spitzer & Semple, 1995; Brand *et al.*, 2002), rabbits (Batra *et al.*, 1997a; Batra *et al.*, 1997b) and free-tailed bat (Grothe & Park, 1998). There is a bias for neurons recorded from the MSO to respond maximally to contralateral leading ITDs. MSO neurons have been denoted EE (to indicate that they are excited in response to both ipsilateral and contralateral sounds).

Sensitivity to ITDs in the envelope of SAM tones has been described from high-CF neurons in the MSO of the cat (Joris, 1996; Yin & Chan, 1990), rabbit (Batra *et al.*, 1997a) and bat (Grothe & Park, 1998). Bats are specialised listeners that use high-frequency sounds echolocation for predation and navigation. All the sensitivity to ITDs described in recordings from the MSO of the bat has been in response to high-frequency SAM tones. Grothe & Park (2000) suggest that the physiological role of the MSO in bats may be as a low-pass filter for detecting amplitude modulated sounds, such as their echolocation calls.

ITD sensitivity in MSO neurons occurs via a process of binaural coincidence detection of stimulus locked inputs. The preservation and enhancement of the temporal properties of ANFs in the pathways that excite the MSO are considered to be a specialisation for the coincidence detection process. The mechanisms underlying sensitivity to ITDs in the MSO and the role of inhibition in the MSO are discussed in Section 1.3.

II.3.7. Lateral Superior Olive (LSO)

Anatomy

The LSO contains bipolar principal neurons arranged tonotopically, with low-frequencies represented in its lateral limb and high-frequencies represented in its medial limb (Tsuchitani & Boudreau, 1966; Guinan *et al.*, 1972; Adam *et al.*, 1999). In contrast to the MSO, the LSO has an over-representation of neurons sensitive to high-frequencies (Guinan *et al.*, 1972). LSO principal neurons receive excitatory input from SBCs in the ipsilateral aVCN and receive inhibitory input from the ipsilateral MNTB (Yin, 2001).

LSO neurons send excitatory projections to both the ipsilateral and contralateral IC, particularly from the high-frequency LSO (Oliver, 2000). LSO neurons send inhibitory projections ipsilaterally to the IC, particularly from the middle to low-frequency LSO. LSO neurons also project bilaterally to the nuclei of the lateral lemniscus (Oliver, 2000).

Physiology

The firing rates of LSO neurons are modulated as a function of IID [as well as the overall intensity and frequency of the stimulus (Tsuchitani & Boudreau, 1969)]. The responses of LSO neurons are predominantly IE; they are inhibited by tones at the

contralateral ear and excited by tones at the ipsilateral ear (Boudreau & Tsuchitani, 1968; Guinan *et al.*, 1972). LSO neurons respond with a characteristic chopping pattern (Zacksenhouse *et al.*, 1995; Adam *et al.*, 1999; Adam *et al.*, 2001). The directional dependence of their firing rates, measured using VAS, is dependent on their modulation in response to IIDs (Tollin & Yin, 2002a; Tollin & Yin, 2002b).

Neurons in the LSO receive well-timed, phase-locked inputs and can also be sensitive to ITDs in response to low-frequency tones (Caird & Klinke, 1983; Finlayson & Caspary, 1991; Batra *et al.*, 1997a; Tollin & Yin, 2005), high-frequency SAM tones (Joris & Yin, 1995; Batra *et al.*, 1997a) and clicks [or transient stimulation (Caird & Klinke, 1983; Wu & Kelly, 1992; Park *et al.*, 1996)]. Sensitivity to ITDs in LSO neurons occurs by a similar coincidence detection process to that which occurs in MSO neurons and is discussed in Section 1.3. Joris and Yin (1998) show that there is a reduction in the upper modulation rate (f_m) at which neurons can phase-lock in the LSO. This may limit the f_m at which sensitivity to ITDs occurs. In the model of peripheral processing employed by Bernstein and Trahiotis (2002), and discussed in Section II.3.3, an additional stage of low-pass filtering is employed to capture the upper limit in the f_m at which sensitivity to high-frequency sounds occurs. This is not necessarily a stage that occurs in the auditory periphery and may reflect processing that occurs in the LSO.

The duplex theory

The MSO has been considered to be specialised for detecting ITDs (Joris *et al.*, 1998) and the LSO has been considered to be specialised for detecting IIDs (Tollin, 2003), therefore providing a physiological correlate of the duplex theory. Neural responses of LSO neurons have been shown to be relatively more modulated in response to changes in IID than ITD, when compared to neurons in the MSO (Caird & Klinke, 1983; Joris & Yin, 1995). Sensitivity to ITDs in the LSO has been considered a by-product of sensitivity to IIDs (Tollin, 2003).

II.3.9. Dorsal Nucleus of the Lateral Lemniscus (DNLL)

Anatomy

The lateral lemniscus is a fibre tract carrying inputs from the SOC and the CN to the IC. The DNLL lies within the lateral lemniscus and receives excitatory input from both the LSO and MSO (Oliver, 2000). The DNLL is tonotopically organized. Neurons in the DNLL send inhibitory, GABAergic (which release γ-amino-butyric-acid) projections to the contralateral IC as well as the contralateral DNLL (Oliver, 2000; Pollak *et al.*, 2002).

Physiology

Neurons in the DNLL show sensitivity to ITDs and to IIDs which is thought to be inherited from the sensitivity of its inputs from the MSO and LSO (Brugge *et al.*, 1970). Inhibition from the DNLL has been implicated in transforming the properties of binaural IC neurons (Pollak *et al.*, 2002; Tollin *et al.*, 2004; D'Angelo *et al.*, 2005; Ingham & McAlpine, 2005).

II.3.10. Inferior Colliculus (IC)

The IC receives inputs from the CN, MSO, LSO and DNLL (Oliver & Huerta, 1992). The IC can be divided into the central nucleus, the dorsal cortex and the paracentral nuclei (Oliver & Huerta, 1992). Neurons in the central nucleus respond to sound with a short latency and are sharply tuned for sound frequency. Therefore, the central nucleus is considered to be, within the IC, the primary neurons of the ascending auditory system. Only the central nucleus of the IC is discussed below.

Anatomy

The IC has a laminar structure with neurons arranged tonotopically along the dorso-ventral axis (low to high-frequency; Schreiner & Langner, 1997). Neurons are either disc-shaped with highly-orientated dendrites (along the frequency-lamina) or are fusiform and probably integrate information across different regions of the IC (Oliver & Huerta, 1992). The IC is considered to be an obligatory relay in the ascending auditory system although evidence does exist for a pathway that may bypass the IC (Malmierca *et al.*, 2002). Inputs from the MSO and LSO may remain non-overlapping, forming synaptic domains with specific binaural properties (Oliver, 2000).

Physiology

Numerous cell types have been identified in the IC based on their electrophysiological properties (Peruzzi *et al.*, 2000). The variety of responses to sound recorded from IC neurons reflects the different pathways which converge on the IC from the CN and SOC (Oliver & Huerta, 1992). The role of the IC in sound localisation has been explored by lesion studies. Unilateral lesions cause deficits in localising sounds within the hemisphere contralateral to the lesion [(Jenkins & Masterton, 1982; Kelly & Kavanagh, 1994); cat and ferret].

Neurons sensitive to ITDs have been widely examined in the IC, partly because *in vivo* recordings from the MSO are difficult to obtain. Classical studies include those by Rose, Yin, Chan, Kuwada and colleagues (Rose *et al.*, 1966; Kuwada & Yin, 1983; Yin & Kuwada, 1983a; Yin & Kuwada, 1983b; Yin *et al.*, 1986; Chan *et al.*, 1987; Yin *et al.*, 1987). Sensitivity to ITDs is believed to be inherited from projections from the MSO and LSO, both directly and indirectly via the DNLL. Several studies have shown that neurons in the IC are sensitive to ITDs in the envelope of high-frequency sounds (Yin *et al.*, 1984; Caird & Klinke, 1987; Batra *et al.*, 1989; Batra *et al.*, 1993; Joris, 2003).

Sensitivity to IIDs has been recorded from neurons in the IC and is believed to be inherited from the LSO as well as created *de novo* in the IC (Sally & Kelly, 1992; Park & Pollak, 1993). High-frequency cells in the IC are often denoted EI (excited by the contralateral ear and inhibited by the ipsilateral ear) which is partly due to contralateral excitation from LSO neurons. IC neurons have been recorded in response to both free-field and VAS stimuli and show a bias towards responding to sounds in the contralateral hemisphere (Semple *et al.*, 1983; Moore *et al.*, 1984; Delgutte *et al.*, 1999; Sterbing *et al.*, 2003).

If IC neurons largely inherit their binaural properties from other brainstem nuclei, it is of interest to ask what processing of spatial information is the IC carrying out. Intracellular recordings reveal that IC neurons receive a mixture of excitatory and inhibitory inputs (Kuwada *et al.*, 1997). It has been suggested that sensitivity to ITDs in the IC might be shaped by the convergence of inputs from cells in the SOC with different tuning to ITDs (McAlpine *et al.*, 1998; Batra *et al.*, 1997b; Shackleton *et al.*, 2000). GABAergic inputs to the IC, possibly from the DNLL, have been studied by iontophoresis of GABA agonists and antagonists in the IC. GABAergic inputs to the IC have been interpreted as sharpening ITD-tuning functions (D'Angelo *et al.*, 2005) or as a gain control mechanism (Ingham & McAlpine, 2005). Inhibitory inputs from the DNLL have also been hypothesised to suppress binaural information subsequent to the onset of a sound, which may be relevant to the precedence effect (Tollin *et al.*, 2004). An emergent property of a small proportion of IC neurons sensitive to ITDs is differential responses to the direction of dynamic ITDs, investigated by changing ITDs with time (Spitzer & Semple, 1991; McAlpine *et al.*, 2000; Ingham *et al.*, 2001). Differences between neural responses to IIDs have been noted between the LSO and the IC. Sensitivity to IIDs occurs at more ipsilateral leading IIDs and there is greater invariance of IID tuning functions with overall intensity in the IC compared to the LSO (Park, 1998; Park *et al.*, 2004). The IC may play a role in extracting sensitivity to the context in which a sound is heard (Malone & Semple, 2001).

II.3.11. Medial Geniculate Nucleus (MGN)

The ventral, medial and lateral divisions of the MGN comprise the auditory thalamus (Winer, 1992). The ventral division of the thalamus has input from the IC and is tonotopically organised (Winer, 1992). Neurons in the ventral division project to primary auditory cortex and have relatively sharp frequency-tuning. Lesions of the thalamus produce deficits in sound localisation (Jenkins & Masterton, 1982). ITD sensitivity in thalamic neurons has been recorded in the rabbit (Stanford *et al.*, 1992). Projections between the thalamus and cortex are diverse and little is known about the functional role of specific pathways for sound localisation.

II.3.12. Auditory Cortex (AC)

The primary auditory cortex (A1) is tonotopically arranged (Winer, 1992). Deficits after unilateral inactivation or lesioning of auditory cortex are primarily in sound localisation in the contralateral hemisphere (Jenkins & Masterton, 1982; Kavanagh & Kelly, 1987; Malhotra *et al.*, 2004). However, even with bilateral lesions of the auditory cortex, Japanese Macaques maintain some ability to discriminate the location of sounds across the midline, although their ability to specifically localise a sound source is severely impaired (Heffner & Heffner, 1990). In humans, patients with right unilateral temporal lobe excisions show greater disruption of sound localisation performance than those with left unilateral temporal lobe excisions, suggesting a right hemispheric specialisation for sound localisation (Zatorre & Penhune, 2001). In contrast to the studies using experimental animals, Zatorre & Penhune (2001) found that the decrease in sound localisation performance was not contralateral to the patients' excision.

Neural firing rates in the auditory cortex are often broadly tuned to spatial position and show a bias towards responding to sounds in the contralateral hemisphere (Stecker *et al.*, 2005b). Classically, responses in the auditory cortex have been

described as onset in nature, although robust firing rates can be obtained e.g. Wang *et al.* (2005).

II.6. Figures and Tables for Appendix II

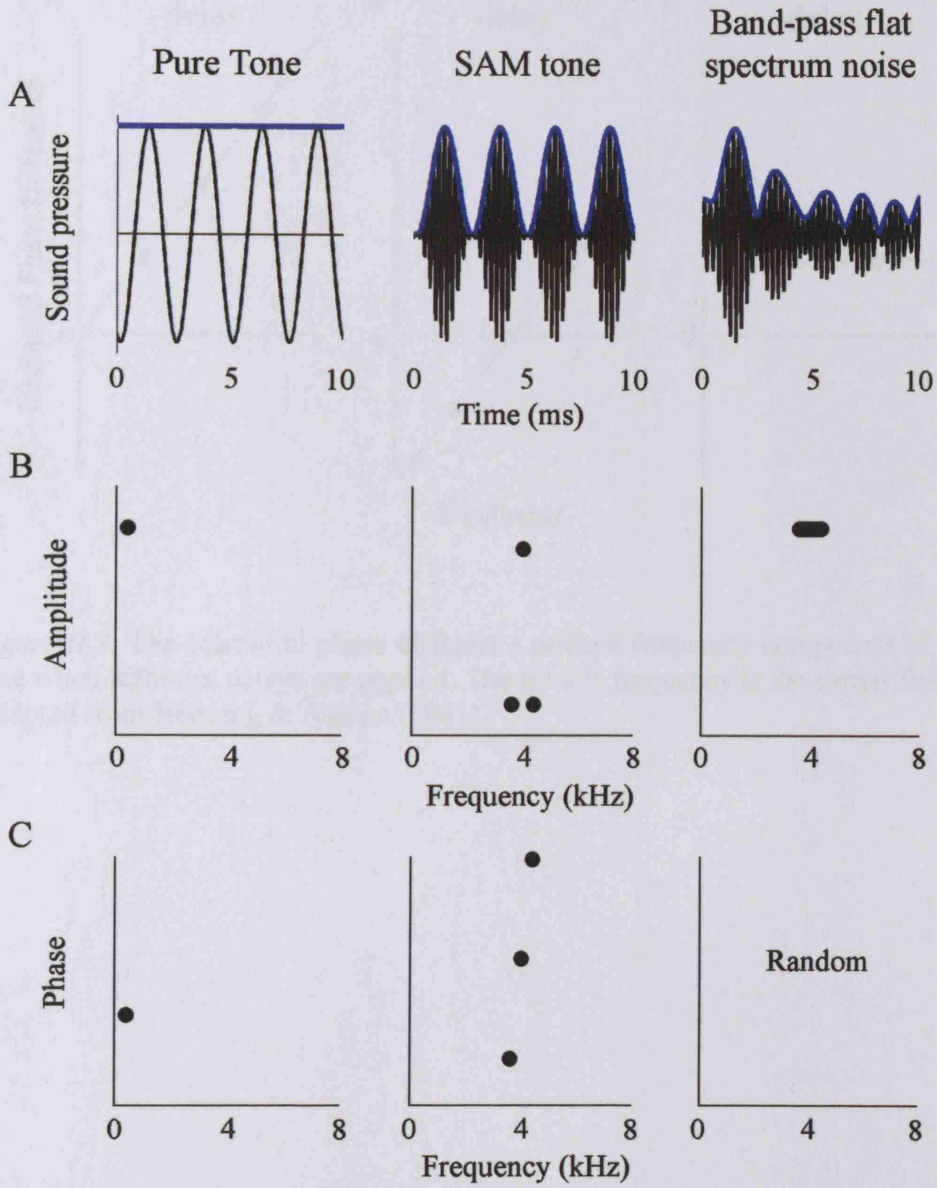

Figure II.1. Diagrams representing a pure tone, a SAM tone and Gaussian noise. A) A portion, in time, of the sound pressure waveforms (black) of a pure tone, a SAM tone and band-pass flat spectrum noise. The thick blue line traces the Hilbert envelope of the sound pressure waveforms. B) The amplitude of the spectrum of the waveforms in A. C) The phase of spectrum of the waveforms in A.

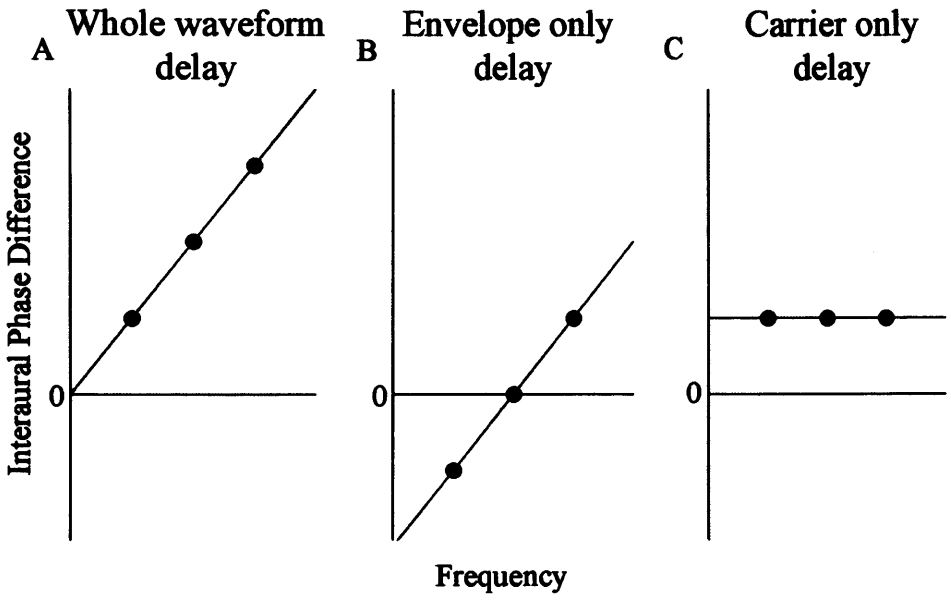

Figure II.2. The interaural phase difference of each frequency component of a SAM tone when different delays are applied. The middle frequency is the carrier frequency. Adapted from Henning & Ashton (1981).

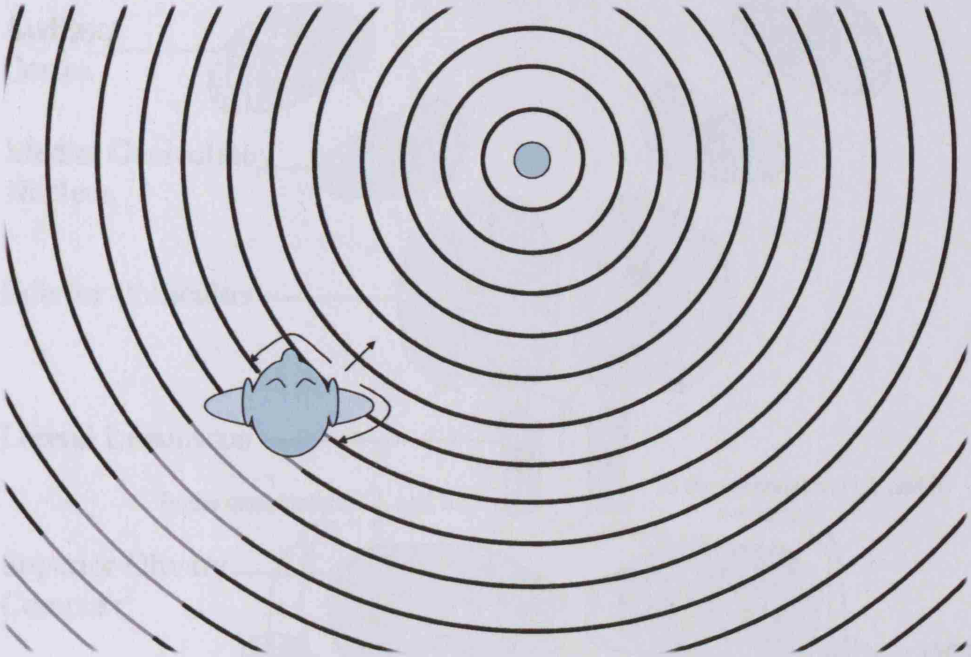

Figure II.3. A diagram (plan view) representing regions of high sound pressure (concentric black circles) emanating from a sound source (blue circle). The head of a human listener (blue!) forms an obstacle to the propagation of the sound and sound is diffracted or reflected (black arrows).

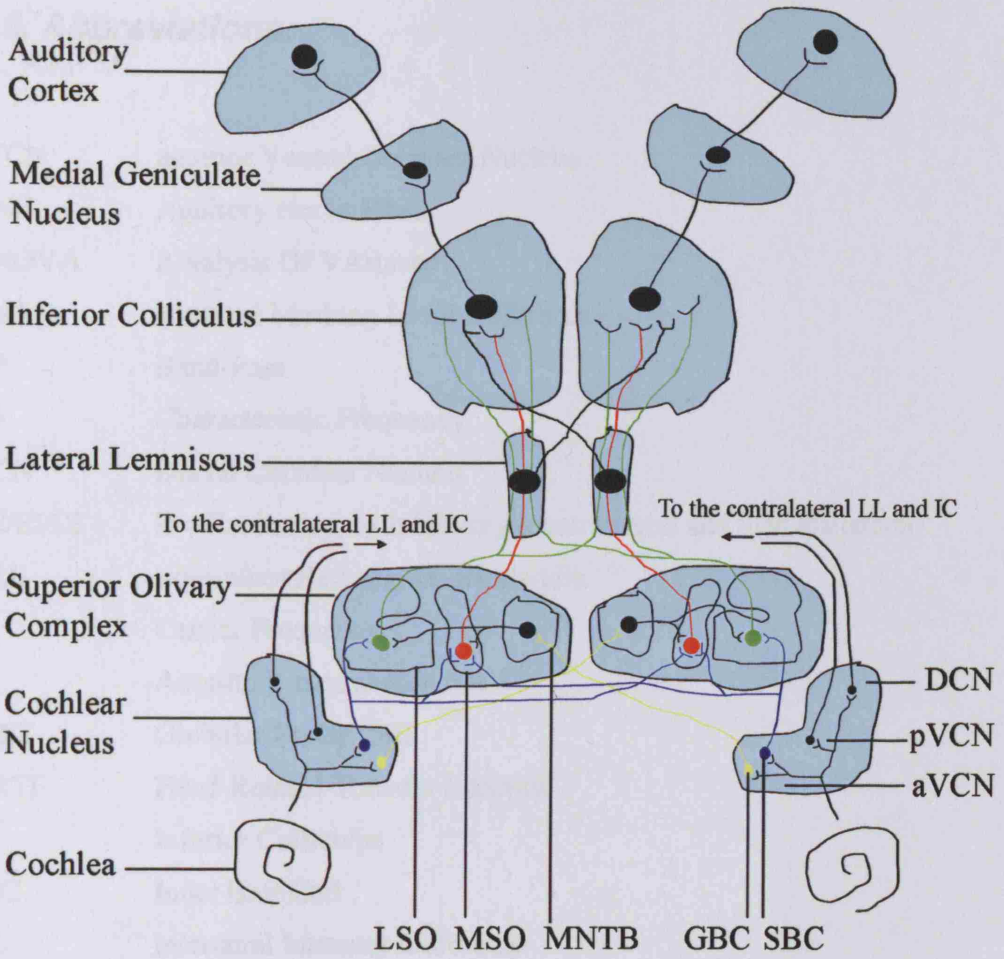

Figure II.4. An overview of the ascending auditory system with emphasis on projections to and from the MSO and LSO. Abbreviations are listed in Section II.5. This diagram is not intended to be fully inclusive of pathways in the auditory system or to be to scale.

II.5. Abbreviations

aVCN	anterior Ventral Cochlear Nucleus
ANF	Auditory Nerve Fibre
ANOVA	ANalysis Of VAriance
BMLD	Binaural Masking Level Difference
BP	Band-Pass
CF	Characteristic Frequency
DCN	Dorsal Cochlear Nucleus
EE/EI/IE	E = Excitatory, I= Inhibitory (contralateral and then ipsilateral)
ERB	Equivalent Rectangular Bandwidth
f_c	Carrier Frequency
f_m	Amplitude modulation rate
GBC	Globular Bushy Cell
HRTF	Head-Related Transfer Function
IC	Inferior Colliculus
IHC	Inner Hair Cell
IID	Interaural Intensity Difference
IPD	Interaural Phase Difference
ITD	Interaural Time Difference
JND	Just Noticeable Difference
LP	Low-Pass
LSO	Lateral Superior Olive
MAA	Minimal Audible Angle
MNTB	Medial Nucleus of the Trapezoid Body
MSO	Medial Superior Olive
OHC	Outer Hair Cell
PSTH	Post Stimulus Time Histogram
pVCN	posterior Ventral Cochlear Nucleus
rMTF	rate Modulation Transfer Function

SAM	Sinusoidally Amplitude Modulated
SBC	Spherical Bushy Cell
SOC	Superior Olivary Complex
sp/s	spikes per second
VAS	Virtual Acoustic Space

II.6. Glossary

ANOVA	Analysis of variance
Anti-phasic	With a phase difference of 0.5 cycles.
Binaural	(Presented to or in response to sound at) the two ears
Broadband transposed noise	A transposed noise created from a low-frequency, flat spectrum noise with low-pass cut-off equal to CF/5 and with carrier frequency equal to CF.
Carrier	For a narrow-band complex sound, the high-frequency waveform (that is modulated in amplitude by an envelope).
Characteristic delay (CD)	The internal delay between the arrival of ipsilateral and contralateral inputs to a coincidence detector neuron.
Characteristic frequency (CF)	The frequency at a neuron's threshold.
Characteristic phase (CP)	The phase difference between to inputs that generates maximal firing.
Closed-field	Sealed with the ear canal.
Contralateral	(on or from) the opposite side.
Correlation	A measure of the similarity between two waveforms.
Cross-correlation	The correlation between two waveforms as a function of the time difference between them.
Dichotic	Different in each ear.
Diotic	The same in both ears.

Envelope	The low-frequency amplitude modulation of a narrow-band waveform. Also equal to the magnitude of the "analytic waveform" representation of a sound.
Fine-structure	The fast fluctuations in a narrow-band waveform (c.f. the envelope).
Fourier series	A series of sinusoids which can describe any waveform. Equal to the spectrum of a waveform..
Free-field	In a non-reverberant environment.
Gating delay	see onset-ITD .
High-CF	A CF of 2 kHz or above.
High-frequency	Above the frequency to which phase-locking to the carrier is possible (taken to be > 2 kHz).
Interaural	Between the ears.
In-phase	With zero phase difference.
Ipsilateral	(on or from) the same side.
ITD-tuning function	The firing rate of an ITD-sensitive neuron as a function of ITD.
Lateralisation	Identification of the location of a sound percept within the head.
Localisation	Identification of the location of a sound percept in space.
Low-CF	A CF of below 2 kHz.
Low-frequency	Below the frequency to which phase-locking to the carrier is no longer possible (taken to be < 2 kHz).
Monaural	(Presented to or in response to sound at) one ear.
Ongoing-ITD	An ITD in the ongoing portion of a waveform.
Onset-ITD	An ITD in the time at which a sound is gated on at each ear.
Peak-ITD (or IPD)	The ITD (or IPD) to which a neuron responds with its maximal firing rate.

Peak-type	An ITD-tuning function in response to a broadband stimulus that is dominated by a peak in firing rate.
Peripheral processing	The effect of the peripheral auditory system (including the basilar membrane, inner hair cells and auditory nerve fibres) on a sound waveform.
Quasi-periodic	Side-peaks and side-troughs are of smaller magnitude than a central peak and/or trough [describing ITD-tuning functions]
Recording	The response to one neuron to SAM or transposed tones at one f_m and over a range of ITDs.
Spectrum	The sinusoidal frequency components which can be added together to form a given waveform. Equal to the Fourier series.
Synchronization rate	The vector strength multiplied by the mean firing rate.
Threshold	The lowest sound intensity to which a neuron responds.
Transposed noise	A low-frequency noise waveform that has been half-wave rectified and multiplied by a carrier of higher frequency.
Transposed tone	A low-frequency tone that has been half-wave rectified and multiplied by a carrier of higher-frequency.
Trough-ITD	The ITD to which a neuron responds with its minimum firing rate.
Trough-type	An ITD-tuning function in response to a broadband stimulus that is dominated by a trough in firing rate.
τ-f plane	A representation of the Jeffress model consisting of a matrix of neurons tuned to different ITDs and different frequencies.
Vector strength	The normalised magnitude of the mean phase vector.

References

Abel, S. M. & Kunov, H. (1983). Lateralization based on interaural phase differences: effects of frequency, amplitude, duration, and shape of rise/decay. *J Acoust Soc Am* **73**, 955-960.

Adam, T. J., Finlayson, P. G., & Schwarz, D. W. (2001). Membrane properties of principal neurons of the lateral superior olive. *J Neurophysiol* **86**, 922-934.

Adam, T. J., Schwarz, D. W., & Finlayson, P. G. (1999). Firing properties of chopper and delay neurons in the lateral superior olive of the rat. *Exp Brain Res* **124**, 489-502.

Agmon-Snir, H., Carr, C. E., & Rinzel, J. (1998). The role of dendrites in auditory coincidence detection. *Nature* **393**, 268-272.

Ashmore, J. F. (1987). A fast motile response in guinea-pig outer hair cells: the cellular basis of the cochlear amplifier. *J Physiol* **388**, 323-347.

Barnes-Davies, M., Barker, M. C., Osmani, F., & Forsythe, I. D. (2004). Kv1 currents mediate a gradient of principal neuron excitability across the tonotopic axis in the rat lateral superior olive. *Eur J Neurosci* **19**, 325-333.

Batra, R., Kuwada, S., & Fitzpatrick, D. C. (1997a). Sensitivity to interaural temporal disparities of low- and high-frequency neurons in the superior olivary complex. I. Heterogeneity of responses. *J Neurophysiol* **78**, 1222-1236.

Batra, R., Kuwada, S., & Fitzpatrick, D. C. (1997b). Sensitivity to interaural temporal disparities of low- and high-frequency neurons in the superior olivary complex. II. Coincidence detection. *J Neurophysiol* **78**, 1237-1247.

Batra, R., Kuwada, S., & Stanford, T. R. (1989). Temporal coding of envelopes and their interaural delays in the inferior colliculus of the unanesthetized rabbit. *J Neurophysiol* **61**, 257-268.

Batra, R., Kuwada, S., & Stanford, T. R. (1993). High-frequency neurons in the inferior colliculus that are sensitive to interaural delays of amplitude-modulated tones: evidence for dual binaural influences. *J Neurophysiol* **70**, 64-80.

Batra, R. & Yin, T. C. (2004). Cross correlation by neurons of the medial superior olive: a reexamination. *J Assoc Res Otolaryngol* **5**, 238-252.

Beckius, G. E., Batra, R., & Oliver, D. L. (1999). Axons from anteroventral cochlear nucleus that terminate in medial superior olive of cat: observations related to delay lines. *J Neurosci* **19**, 3146-3161.

Bernstein, L. R. & Trahiotis, C. (1982). Detection of interaural delay in high-frequency noise. *J Acoust Soc Am* **71**, 147-152.

Bernstein, L. R. & Trahiotis, C. (1985a). Lateralization of low-frequency, complex waveforms: the use of envelope-based temporal disparities. *J Acoust Soc Am* **77**, 1868-1880.

Bernstein, L. R. & Trahiotis, C. (1985b). Lateralization of sinusoidally amplitude-modulated tones: effects of spectral locus and temporal variation. *J Acoust Soc Am* **78**, 514-523.

Bernstein, L. R. & Trahiotis, C. (1992). Discrimination of interaural envelope correlation and its relation to binaural unmasking at high frequencies. *J Acoust Soc Am* **91**, 306-316.

Bernstein, L. R. & Trahiotis, C. (1994). Detection of interaural delay in high-frequency sinusoidally amplitude-modulated tones, two-tone complexes, and bands of noise. *J Acoust Soc Am* **95**, 3561-3567.

Bernstein, L. R. & Trahiotis, C. (1996a). On the use of the normalized correlation as an index of interaural envelope correlation. *J Acoust Soc Am* **100**, 1754-1763.

Bernstein, L. R. & Trahiotis, C. (1996b). The normalized correlation: accounting for binaural detection across center frequency. *J Acoust Soc Am* **100**, 3774-3784.

Bernstein, L. R. & Trahiotis, C. (2002). Enhancing sensitivity to interaural delays at high frequencies by using "transposed stimuli". *J Acoust Soc Am* **112**, 1026-1036.

Bernstein, L. R. & Trahiotis, C. (2003). Enhancing interaural-delay-based extents of laterality at high frequencies by using "transposed stimuli". *J Acoust Soc Am.* **113**, 3335-3347.

Best, V., van Schaik, A., & Carlile, S. (2004). Separation of concurrent broadband sound sources by human listeners. *J Acoust Soc Am* **115**, 324-336.

Blauert, J. (1997). *Spatial Hearing: The Psychophysics of Human Sound Localization* Cambridge, Massachussetts: MIT Press.

Boudreau, J. C. & Tsuchitani, C. (1968). Binaural interaction in the cat superior olive S segment. *J Neurophysiol* **31**, 442-454.

Brand, A., Behrend, O., Marquardt, T., McAlpine, D., & Grothe, B. (2002). Precise inhibition is essential for microsecond interaural time difference coding. *Nature* **417**, 543-547.

Brew, H. M. & Forsythe, I. D. (1995). Two voltage-dependent K+ conductances with complementary functions in postsynaptic integration at a central auditory synapse. *J Neurosci* **15**, 8011-8022.

Brew, H. M. & Forsythe, I. D. (2005). Systematic variation of potassium current amplitudes across the tonotopic axis of the rat medial nucleus of the trapezoid body. *Hear Res* **206**, 116-132.

Brugge, J. F., Anderson, D. J., & Aitkin, L. M. (1970). Responses of neurons in the dorsal nucleus of the lateral lemniscus of cat to binaural tonal stimulation. *J Neurophysiol* **33**, 441-458.

Brugge, J. F., Anderson, D. J., Hind, J. E., & Rose, J. E. (1969). Time structure of discharges in single auditory nerve fibers of the squirrel monkey in response to complex periodic sounds. *J Neurophysiol* **32**, 386-401.

Brugge, J.F., Reale, R.A., & Hind, J.E. (1996). The structure of spatial receptive fields of neurons in primary auditory cortex of the cat. *J Neurosci* **16**, 4420-4437

Buell, T. N. & Hafter, E. R. (1988). Discrimination of interaural differences of time in the envelopes of high-frequency signals: integration times. *J Acoust Soc Am* **84**, 2063-2066.

Buell, T. N., Trahiotis, C., & Bernstein, L. R. (1991). Lateralization of low-frequency tones: relative potency of gating and ongoing interaural delays. *J Acoust Soc Am* **90**, 3077-3085.

Bullock, D. C., Palmer, A. R., & Rees, A. (1988). Compact and easy-to-use tungsten-in-glass microelectrode manufacturing workstation. *Med Biol Eng Comput* **26**, 669-672.

Caird, D. & Klinke, R. (1983). Processing of binaural stimuli by cat superior olivary complex neurons. *Exp Brain Res* **52**, 385-399.

Caird, D. & Klinke, R. (1987). Processing of interaural time and intensity differences in the cat inferior colliculus. *Exp Brain Res* **68**, 379-392.

Campbell, R. A., Doubell, T. P., Nodal, F. R., Schnupp, J. W., & King, A. J. (2005). Interaural Timing Cues do not Contribute to the Map of Space in the Ferret Superior Colliculus: a Virtual Acoustic Space Study. *J Neurophysiol* **95**, 242-254

Cant, N. B. (1992). The Cochlear nucleus: Neuronal Types and Their Synaptic Organization. In *The Mammalian Auditory Pathway: Neuroanatomy*, eds. Webster, D. B., Popper, A. N., & Fay, R. R., pp. 66-116. Springer-Verlag, New York.

Cant, N. B. & Benson, C. G. (2003). Parallel auditory pathways: projection patterns of the different neuronal populations in the dorsal and ventral cochlear nuclei. *Brain Res Bull* **60**, 457-474.

Carr, C. E. & Konishi, M. (1988). Axonal delay lines for time measurement in the owl's brainstem. *Proc Natl Acad Sci U.S.A* **85**, 8311-8315.

Carr, C. E. & Konishi, M. (1990). A circuit for detection of interaural time differences in the brain stem of the barn owl. *J Neurosci* **10**, 3227-3246.

Chan, J. C., Yin, T. C., & Musicant, A. D. (1987). Effects of interaural time delays of noise stimuli on low-frequency cells in the cat's inferior colliculus. II. Responses to band-pass filtered noises. *J Neurophysiol* **58**, 543-561.

Clark, G. (2003). *Cochlear Implants: fundamentals and applications* Springer Science + Business media, Inc., New York.

Colburn, H. S. (1973). Theory of binaural interaction based on auditory-nerve data. I. General strategy and preliminary results on interaural discrimination. *J Acoust Soc Am* **54**, 1458-1470.

Colburn, H. S. & Esquissaud, P. (1976). An auditory-nerve model for interaural time discrimination of high-frequency complex stimuli. *J Acoust Soc Am* Suppl.1 59 [S23].

Crow, G., Langford, T.L., & Moushegian, G. (1980) Coding of interaural time differences by some high-frequency neurons of the inferior colliculus: Responses to noise bands and two-tone complexes. *Hear Res* **3**, 147-153

Crow, G., Rupert, A. L., & Moushegian, G. (1978). Phase locking in monaural and binaural medullary neurons: implications for binaural phenomena. *J Acoust Soc Am* **64**, 493-501.

D'Angelo, W. R., Sterbing, S. J., Ostapoff, E. M., & Kuwada, S. (2005). Role of GABAergic inhibition in the coding of interaural time differences of low-frequency sounds in the inferior colliculus. *J Neurophysiol* **93**, 3390-3400.

David, E. E., Guttman, N., & van Bergeijk, W. A. (1959). Binaural interaction of high-frequency complex stimuli. *J Acoust Soc Am* **31**, 774-782.

Delgutte, B., Joris, P. X., Litovsky, R. Y., & Yin, T. C. (1999). Receptive fields and binaural interactions for virtual-space stimuli in the cat inferior colliculus. *J Neurophysiol* **81**, 2833-2851.

Dodson, P. D., Barker, M. C., & Forsythe, I. D. (2002). Two heteromeric Kv1 potassium channels differentially regulate action potential firing. *J Neurosci* **22**, 6953-6961.

Domnitz, R. H. & Colburn, H.S. (1977). Lateral position and interaural discrimination. *J Acoust Soc Am* **61**, 1586-1598.

Dreyer, A., Oxenham, A., & Delgutte, B. (2006). Phase locking of auditory-nerve fibers to the envelopes of high-frequency sounds: Implications for sound localization. *J Neurophysiol* Epub ahead of print, Jun 28

Endler, F., Farkas, D., Asadollahi, A., Nelken, I., & Wagner, H. (2005). Correlates of Binaural Masking Level Difference measured in High Frequency Neurons in Cats and Owls. Assoc. Res. Otolaryngol. Abs.: 655.

Evans, E. F., Pratt, S. R., Spenner, H., & Cooper, N. P. (1992). Comparisons of Physiological and Behavioural Properties: Auditory Frequency Selectivity. *Advances in the Biosciences* **83**, 159-164.

Ewert, S. D. & Dau, T. (2000). Characterizing frequency selectivity for envelope fluctuations. *J Acoust Soc Am* **108**, 1181-1196.

Fay, R. R. (1988). *Hearing in Vertebrates: a Psychophysics Databook* Hill-Fay Associates, Winnetka, Illinois.

Fettiplace, R. & Fuchs, P. A. (1999). Mechanisms of hair cell tuning. *Annu Rev Physiol* **61**, 809-834.

Finlayson, P. G. & Caspary, D. M. (1991). Low-frequency neurons in the lateral superior olive exhibit phase-sensitive binaural inhibition. *J Neurophysiol* **65**, 598-605.

Fitzpatrick, D. C., Batra, R., Stanford, T. R., & Kuwada, S. (1997). A neuronal population code for sound localization. *Nature* **388**, 871-874.

Fitzpatrick, D. C., Kuwada, S., & Batra, R. (2000). Neural sensitivity to interaural time differences: beyond the Jeffress model. *J Neurosci* **20**, 1605-1615.

Fitzpatrick, D. C., Kuwada, S., & Batra, R. (2002). Transformations in processing interaural time differences between the superior olivary complex and inferior colliculus: beyond the Jeffress model. *Hear Res* **168**, 79-89.

Forsythe, I. D. (1994). Direct patch recording from identified presynaptic terminals mediating glutamatergic EPSCs in the rat CNS, in vitro. *J Physiol* **479 (Pt 3)**, 381-387.

Forsythe, I. D. & Barnes-Davies, M. (1993). The binaural auditory pathway: excitatory amino acid receptors mediate dual timecourse excitatory postsynaptic currents in the rat medial nucleus of the trapezoid body. *Proc Biol Sci* **251**, 151-157.

Freyman, R. L., Clifton, R. K., & Litovsky, R. Y. (1991). Dynamic processes in the precedence effect. *J Acoust Soc Am* **90**, 874-884.

Freyman, R. L., Zurek, P. M., Balakrishnan, U., & Chiang, Y. C. (1997). Onset dominance in lateralization. *J Acoust Soc Am* **101**, 1649-1659.

Frisina, R. D. (2001). Subcortical neural coding mechanisms for auditory temporal processing. *Hear Res* **158**, 1-27.

Furukawa, S. & Middlebrooks, J. C. (2002). Cortical representation of auditory space: information-bearing features of spike patterns. *J Neurophysiol* **87**, 1749-1762.

Gaik, W. (1993). Combined evaluation of interaural time and intensity differences: psychoacoustic results and computer modeling. *J Acoust Soc Am* **94**, 98-110.

Galambos, R., Schwartzkopff, J., & Rupert, A. L. (1959). Microelectrode study of superior olivary nuclei. *J Neurophysiol* **197**, 527-536.

Gardner, M. B. (1968). Historical background of the Haas and-or precedence effect. *J Acoust Soc Am* **43**, 1243-1248.

Gardner, M. B. & Gardner, R. S. (1973). Problem of localization in the median plane: effect of pinnae cavity occlusion. *J Acoust Soc Am* **53**, 400-408.

Geisler, C. D., Rhode, W. S., & Hazelton, D. W. (1969). Responses of inferior colliculus neurons in the cat to binaural acoustic stimuli having wide-band spectra. *J Neurophysiol* **32**, 960-974.

Goldberg, J. M. & Brown, P. B. (1969). Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. *J Neurophysiol* **32**, 613-636.

Golding, N. L., Robertson, D., & Oertel, D. (1995). Recordings from slices indicate that octopus cells of the cochlear nucleus detect coincident firing of auditory nerve fibers with temporal precision. *J Neurosci* **15**, 3138-3153.

Grantham, D.W. (1984) Interaural intensity discrimination: Insensitivity at 1000 Hz. *J Acoust Soc Am* **75**, 1191-1194.

Griffin, S. J., Bernstein, L. R., Ingham, N. J., & McAlpine, D. (2005). Neural sensitivity to interaural envelope delays in the inferior colliculus of the guinea pig. *J Neurophysiol* **93**, 3463-3478.

Grothe, B. & Park, T. J. (1998). Sensitivity to interaural time differences in the medial superior olive of a small mammal, the Mexican free-tailed bat. *J Neurosci* **18**, 6608-6622.

Grothe, B. & Park, T. J. (2000). Structure and function of the bat superior olivary complex. *Microsc Res Tech* **51**, 382-402.

Grothe, B., Park, T. J., & Schuller, G. (1997). Medial superior olive in the free-tailed bat: response to pure tones and amplitude-modulated tones. *J Neurophysiol* **77**, 1553-1565.

Grothe, B. & Sanes, D. H. (1993). Bilateral inhibition by glycinergic afferents in the medial superior olive. *J Neurophysiol* **69**, 1192-1196.

Grothe, B. & Sanes, D. H. (1994). Synaptic inhibition influences the temporal coding properties of medial superior olivary neurons: an in vitro study. *J Neurosci* **14**, 1701-1709.

Guinan, J. J. Jr., Norris, B. E., & Guinan, S. S. (1972). Single auditory units in the superior olivary complex. II. Locations of unit categories and tonotopic organization. *Int J Neurosci* **4**, 147-166.

Hafter, E. R., Bourbon, W. T., Blocker, A. S., & Tucker, A. (1969). A direct comparison between lateralization and detection under conditions of antiphasic masking. *J Acoust Soc Am* **46**, 1452-1457.

Hafter, E. R. & Buell, T. N. (1990). Restarting the adapted binaural system. *J Acoust Soc Am* **88**, 806-812.

Hafter, E. R. & Carrier, S. C. (1972). Binaural interaction in low-frequency stimuli: the inability to trade time and intensity completely. *J Acoust Soc Am* **51**, 1852-1862.

Hafter, E. R., De Maio, J., & Hellman, W. S. (1975). Difference thresholds for interaural delay. *J Acoust Soc Am* **57**, 181-187.

Hafter, E. R., Dye, R. H., Jr., & Gilkey, R. H. (1979). Lateralization of tonal signals which have neither onsets nor offsets. *J Acoust Soc Am* **65**, 471-477.

Hafter, E. R. & Jeffress, L. A. (1968). Two-image lateralization of tones and clicks. *J Acoust Soc Am* **44**, 563-569.

Hancock, K. E. & Delgutte, B. (2004). A physiologically based model of interaural time difference discrimination. *J Neurosci* **24**, 7110-7117.

Harper, N. S. & McAlpine, D. (2004). Optimal neural population coding of an auditory spatial cue. *Nature* **430**, 682-686.

Harris, G. G. (1960). Binaural interactions of impulsive stimuli and pure tones. *J Acoust Soc Am* **32**, 685-692.

Hartmann, W. M. (1983). Localization of sound in rooms. *J Acoust Soc Am* **74**, 1380-1391.

Hartman,W. M. (1998). Signals, Sound, and Sensation. Springer-Verlag, New York.

Hartmann, W. M. & Wittenberg, A. (1996). On the externalization of sound images. *J Acoust Soc Am* **99**, 3678-3688.

Hartung, K. & Trahiotis, C. (2001). Peripheral auditory processing and investigations of the "precedence effect" which utilize successive transient stimuli. *J Acoust Soc Am* **110**, 1505-1513.

Heffner, H. E. & Heffner, R. S. (1990). Effect of bilateral auditory cortex lesions on sound localization in Japanese macaques. *J Neurophysiol* **64**, 915-931.

Heffner, H. E. & Heffner, R. S. (2005). The sound-localization ability of cats. *J Neurophysiol* **94**, 3653-3655.

Heffner, R. S. & Heffner, H. E. (1988). Sound localization acuity in the cat: effect of azimuth, signal duration, and test procedure. *Hear Res* **36**, 221-232.

Heffner, R. S. & Heffner, H. E. (1992). Visual factors in sound localization in mammals. *J Comp Neurol* **317**, 219-232.

Heil, P. & Neubauer, H. (2001). Temporal integration of sound pressure determines thresholds of auditory-nerve fibers. *J Neurosci* **21**, 7404-7415.

Heil, P. & Neubauer, H. (2003). A unifying basis of auditory thresholds based on temporal summation. *Proc Natl Acad Sci U.S.A* **100**, 6151-6156.

Henning, G. B. (1974a). Detectability of interaural delay in high-frequency complex waveforms. *J Acoust Soc Am* **55**, 84-90.

Henning, G. B. (1974b). Lateralization and the binaural masking-level difference. *J Acoust Soc Am* **55**, 1259-1262.

Henning, G. B. (1980). Some observations on the lateralization of complex waveforms. *J Acoust Soc Am* **68**, 446-454.

Henning, G. B. & Ashton, J. (1981). The effect of carrier and modulation frequency on lateralization based on interaural phase and interaural group delay. *Hear Res* **4**, 185-194.

Hirsh, I. (1948). The influence of interaural phase on interaural summation and inhibition. *J Acoust Soc Am* **20**, 536-544.

Houtgast, T. & Aoki, S. (1994). Stimulus-onset dominance in the perception of binaural information. *Hear Res* **72**, 29-36.

Houtgast, T. & Plomp, R. (1968). Lateralization threshold of a signal in noise. *J Acoust Soc Am* **44**, 807-812.

Hudspeth, A. J. (2005). How the ear's works work: mechanoelectrical transduction and amplification by hair cells. *C R Biol* **328**, 155-162.

Ingham, N. J., Hart, H. C., & McAlpine, D. (2001). Spatial receptive fields of inferior colliculus neurons to auditory apparent motion in free field. *J Neurophysiol* **85**, 23-33.

Ingham, N. J. & McAlpine, D. (2005). GABAergic inhibition controls neural gain in inferior colliculus neurons sensitive to interaural time differences. *J Neurosci* **25**, 6187-6198.

Jeffress, L. A. (1948). A place theory of sound localization. *J Comp Physiol Psych* 61, 468-486.

Jeffress, L. A., Blodgett, H. C., & Deatherage, B. H. (1952). The masking of tones by white noise as a function of the interaural phases of both components. I. 500 cycles. *J Acoust Soc Am* **24**, 523-527.

Jeffress, L. A., Blodgett, H. C., & Deatherage, B. H. (1962). Masking and interaural phase. II. 167 cycles. *J Acoust Soc Am* **34**, 1124-1126.

Jenkins, W. M. & Masterton, R. B. (1982). Sound localization: effects of unilateral lesions in central auditory system. *J Neurophysiol* **47**, 987-1016.

Johnson, D. H. (1980). The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. *J Acoust Soc Am* **68**, 1115-1122.

Jones, P. J. & Williams, R. P. (1981). An experiment to determine whether the interaural time differences used in lateralizing middle and high-frequency complex tones is dependent in any way on fine structure information. *Acustica* **47**, 164-169.

Jonson, K. M., Lyle, J. G., Edwards, M. J., & Penny, R. H. C. (1975). Problems in behavioural research with the guinea pig: a selective review. *Anim Behav* **23**, 632-639.

Joris, P. X. (1996). Envelope coding in the lateral superior olive. II. Characteristic delays and comparison with responses in the medial superior olive. *J Neurophysiol* **76**, 2137-2156.

Joris, P. X. (2003). Interaural time sensitivity dominated by cochlea-induced envelope patterns. *J Neurosci* **23**, 6345-6350.

Joris, P. X., Carney, L. H., Smith, P. H., & Yin, T. C. (1994a). Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency. *J Neurophysiol* **71**, 1022-1036.

Joris, P. X., Smith, P. H., & Yin, T. C. (1994b). Enhancement of neural synchronization in the anteroventral cochlear nucleus. II. Responses in the tuning curve tail. *J Neurophysiol* **71**, 1037-1051.

Joris, P. X., Smith, P. H., & Yin, T. C. (1998). Coincidence detection in the auditory system: 50 years after Jeffress. *Neuron* **21**, 1235-1238.

Joris, P. X., Van De, S. B., & van der, H. M. (2005). Temporal damping in response to broadband noise. I. Inferior colliculus. *J Neurophysiol* **93**, 1857-1870.

Joris, P. X. & Yin, T. C. (1992). Responses to amplitude-modulated tones in the auditory nerve of the cat. *J Acoust Soc Am* **91**, 215-232.

Joris, P. X. & Yin, T. C. (1995). Envelope coding in the lateral superior olive. I. Sensitivity to interaural time differences. *J Neurophysiol* **73**, 1043-1062.

Joris, P. X. & Yin, T. C. (1998). Envelope coding in the lateral superior olive. III. Comparison with afferent pathways. *J Neurophysiol* **79**, 253-269.

Kapfer, C., Seidl, A. H., Schweizer, H., & Grothe, B. (2002). Experience-dependent refinement of inhibitory inputs to auditory coincidence-detector neurons. *Nat Neurosci* **5**, 247-253.

Kavanagh, G. L. & Kelly, J. B. (1987). Contribution of auditory cortex to sound localization by the ferret (Mustela putorius). *J Neurophysiol* **57**, 1746-1766.

Kavanagh, G. L. & Kelly, J. B. (1992). Midline and lateral field sound localization in the ferret (Mustela putorius): contribution of the superior olivary complex. *J Neurophysiol* **67**, 1643-1658.

Kelly, J. B. & Kavanagh, G. L. (1994). Sound localization after unilateral lesions of inferior colliculus in the ferret (Mustela putorius). *J Neurophysiol* **71**, 1078-1087.

Kemp, D. T. (1978). Stimulated acoustic emissions from within the human auditory system. *J Acoust Soc Am* **64**, 1386-1391.

Kennedy, H. J., Crawford, A. C., & Fettiplace, R. (2005). Force generation by mammalian hair bundles supports a role in cochlear amplification. *Nature* **433**, 880-883.

Klumpp, R. G. & Eady, H. R. (1956). Some measurements of interaural time difference thresholds. *J Acoust Soc Am* **28**, 859-860.

Knudsen, E. I. & Konishi, M. (1978). A neural map of auditory space in the owl. *Science* **200**, 795-797.

Kohlrausch, A., Fassel, R., & Dau, T. (2000). The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers. *J Acoust Soc Am* **108**, 723-734.

Kopp-Scheinpflug, C., Lippe, W. R., Dorrscheidt, G. J., & Rubsamen, R. (2003). The medial nucleus of the trapezoid body in the gerbil is more than a relay: comparison of pre- and postsynaptic activity. *J Assoc Res Otolaryngol* **4**, 1-23.

Krishna, B. S. & Semple, M. N. (2000). Auditory temporal processing: responses to sinusoidally amplitude-modulated tones in the inferior colliculus. *J Neurophysiol* **84**, 255-273.

Kuhn, G. F. (1977). Model for the interaural time differences in the azimuthal plane. *J Acoust Soc Am* **62**, 157-167.

Kulkarni, A. & Colburn, H. S. (1998). Role of spectral detail in sound-source localization. *Nature* **396**, 747-749.

Kuwada, S., Batra, R., Yin, T. C., Oliver, D. L., Haberly, L. B., & Stanford, T. R. (1997). Intracellular recordings in response to monaural and binaural stimulation of neurons in the inferior colliculus of the cat. *J Neurosci* **17**, 7565-7581.

Kuwada, S., Stanford, T. R., & Batra, R. (1987). Interaural phase-sensitive units in the inferior colliculus of the unanesthetized rabbit: effects of changing frequency. *J Neurophysiol* **57**, 1338-1360.

Kuwada, S. & Yin, T. C. (1983). Binaural interaction in low-frequency neurons in inferior colliculus of the cat. I. Effects of long interaural delays, intensity, and repetition rate on interaural delay function. *J Neurophysiol* **50**, 981-999.

Kuwada, S., Yin, T. C., & Wickesberg, R. E. (1979). Response of cat inferior colliculus neurons to binaural beat stimuli: possible mechanisms for sound localization. *Science* **206**, 586-588.

Laback, B., Pok, S. M., Baumgartner, W. D., Deutsch, W. A., & Schmid, K. (2004). Sensitivity to interaural level and envelope time differences of two bilateral cochlear implant listeners using clinical sound processors. *Ear Hear* **25**, 488-500.

Langford, T. I. & Jeffress, L. A. (1964). Effect of noise cross correlation on binaural signal detection. *J Acoust Soc Am* **36**, 1455-1458.

Leakey, D. M., Sayers, B. M., & Cherry, C. (1958). Binaural fusion of low- and high-frequency sounds. *J Acoust Soc Am* **30**, 222.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *J Acoust Soc Am* **49**, Suppl.

Licklider, J. C. R. (1948). The influence of interaural phase relations upon the masking of speech by white noise. *J Acoust Soc Am* **20**, 150-159.

Lin, T. & Guinan, J. J., Jr. (2000). Auditory-nerve-fiber responses to high-level clicks: interference patterns indicate that excitation is due to the combination of multiple drives. *J Acoust Soc Am* **107**, 2615-2630.

Litovsky, R. Y., Colburn, H. S., Yost, W. A., & Guzman, S. J. (1999). The precedence effect. *J Acoust Soc Am* **106**, 1633-1654.

Loftus, W. C., Bishop, D. C., Saint Marie, R. L., & Oliver, D. L. (2004). Organization of binaural excitatory and inhibitory inputs to the inferior colliculus from the superior olive. *J Comp Neurol* **472**, 330-344.

Louage, D. H., van der, H. M., & Joris, P. X. (2004). Temporal properties of responses to broadband noise in the auditory nerve. *J Neurophysiol* **91**, 2051-2065.

Louage, D. H., van der, H. M., & Joris, P. X. (2005). Enhanced temporal response properties of anteroventral cochlear nucleus neurons to broadband noise. *J Neurosci* **25**, 1560-1570.

Macpherson, E. A. & Middlebrooks, J. C. (2002). Listener weighting of cues for lateral angle: the duplex theory of sound localization revisited. *J Acoust Soc Am* **111**, 2219-2236.

Magnusson, A. K., Kapfer, C., Grothe, B., & Koch, U. (2005). Maturation of Glycinergic Inhibition in the Gerbil Medial Superior Olive after Hearing Onset. *J Physiol* **568**, 497-512.

Maier, J. & Klump, G. M. (2005). Directional Hearing in the Gerbil (Meriones Unguiculatus). *J Acoust Soc Am* **119**, 1029-1036.

Malhotra, S., Hall, A. J., & Lomber, S. G. (2004). Cortical control of sound localization in the cat: unilateral cooling deactivation of 19 cerebral areas. *J Neurophysiol* **92**, 1625-1643.

Malmierca, M. S., Merchan, M. A., Henkel, C. K., & Oliver, D. L. (2002). Direct projections from cochlear nuclear complex to auditory thalamus in the rat. *J Neurosci* **22**, 10891-10897.

Malone, B. J. & Semple, M. N. (2001). Effects of auditory stimulus context on the representation of frequency in the gerbil inferior colliculus. *J Neurophysiol* **86**, 1113-1130.

Mardia, K. V. & Jupp, P. E. (1999). *Directional Statistics.* John Wiley & Sons.

Marquardt, T. & McAlpine, D. (2001). Simulation of binaural unmasking just using four binaural channels. Assoc. Res. Otolaryngol. Abs.: 21716.

Martin, R. L., Paterson, M., & McAnally, K. I. (2004). Utility of monaural spectral cues is enhanced in the presence of cues to sound-source lateral angle. *J Assoc Res Otolaryngol* **5**, 80-89.

McAlpine, D., Jiang, D., & Palmer, A. R. (1996). Interaural delay sensitivity and the classification of low best-frequency binaural responses in the inferior colliculus of the guinea pig. *Hear Res* **97**, 136-152.

McAlpine, D., Jiang, D., & Palmer, A. R. (2001). A neural code for low-frequency sound localization in mammals. *Nat Neurosci* **4**, 396-401.

McAlpine, D., Jiang, D., Shackleton, T. M., & Palmer, A. R. (1998). Convergent input from brainstem coincidence detectors onto delay-sensitive neurons in the inferior colliculus. *J Neurosci* **18**, 6026-6039.

McAlpine, D., Jiang, D., Shackleton, T. M., & Palmer, A. R. (2000). Responses of neurons in the inferior colliculus to dynamic interaural phase cues: evidence for a mechanism of binaural adaptation. *J Neurophysiol* **83**, 1356-1365.

McFadden, D. & Pasanen, E. G. (1976). Lateralization of high frequencies based on interaural time differences. *J Acoust Soc Am* **59**, 634-639.

McFadden, D. & Pasanen, E. G. (1978). Binaural detection at high frequencies with time-delayed waveforms. *J Acoust Soc Am* **63**, 1120-1131.

Medvedeva, M. V. (1977). Stereotaxic coordinates of midbrain structures of the Guinea Pig. *Fiziologicheskii Zhurnal SSSR imeni I M Sechenova* **63**, 756-759.

Middlebrooks, J. C. & Green, D. M. (1990). Directional dependence of interaural envelope delays. *J Acoust Soc Am* **87**, 2149-2162.

Middlebrooks, J. C. & Green, D. M. (1991). Sound localization by human listeners. *Annu Rev Psychol* **42**, 135-159.

Middlebrooks, J. C. & Knudsen, E. I. (1984). A neural code for auditory space in the cat's superior colliculus. *J Neurosci* **4**, 2621-2634.

Middlebrooks, J. C., Makous, J. C., & Green, D. M. (1989). Directional sensitivity of sound-pressure levels in the human ear canal. *J Acoust Soc Am* **86**, 89-108.

Mills, A. W. (1958). On the minimum audible angle. *J Acoust Soc Am* **30**, 237-246.

Mills, A. W. (1960). Lateralization of high-frequency tones. *J Acoust Soc Am* 32, 132-134.

Moore, B. C. J. (1997). Frequency analysis and pitch perception. pp. 1447-1460. John Wiley and Sons, New York.

Moore, B. C. J. (2004). *An Introduction to the Psychology of Hearing*, 5 ed. Elsevier Academic Press, London.

Moore, D. R., Semple, M. N., Addison, P. D., & Aitkin, L. M. (1984). Properties of spatial receptive fields in the central nucleus of the cat inferior colliculus. I. Responses to tones of low intensity. *Hear Res* **13**, 159-174.

Moushegian, G., Rupert, A. L., & Gidda, J. S. (1975). Functional characteristics of superior olivary neurons to binaural stimuli. *J Neurophysiol* **38**, 1037-1048.

Muller, M., Robertson, D., & Yates, G. K. (1991). Rate-versus-level functions of primary auditory nerve fibres: evidence for square law behaviour of all fibre categories in the guinea pig. *Hear Res* **55**, 50-56.

Musicant, A. D., Chan, J. C., & Hind, J. E. (1990). Direction-dependent spectral properties of cat external ear: new data and cross-species comparisons. *J Acoust Soc Am* **87**, 757-781.

Nelken, I., Chechik, G., Mrsic-Flogel, T. D., King, A. J., & Schnupp, J. W. (2005). Encoding stimulus information by spike numbers and mean response time in primary auditory cortex. *J Comput Neurosci* **19**, 199-221.

Nuetzel, J. M. & Hafter, E. R. (1976). Lateralization of complex waveforms: effects of fine structure, amplitude, and duration. *J Acoust Soc Am* **60**, 1339-1346.

Nuetzel, J. M. & Hafter, E. R. (1981). Discrimination of interaural delays in complex waveforms: Spectral effects. *J Acoust Soc Am* **69**, 1112-1118.

Oertel, D. (1999). The role of timing in the brain stem auditory nuclei of vertebrates. *Annu Rev Physiol* **61**, 497-519.

Oertel, D., Bal, R., Gardner, S. M., Smith, P. H., & Joris, P. X. (2000). Detection of synchrony in the activity of auditory nerve fibers by octopus cells of the mammalian cochlear nucleus. *Proc Natl Acad Sci U.S.A* **97**, 11773-11779.

Oliver, D. L. (2000). Ascending efferent projections of the superior olivary complex. *Microsc Res Tech* **51**, 355-363.

Oliver, D. L., Beckius, G. E., Bishop, D. C., Loftus, W. C., & Batra, R. (2003). Topography of interaural temporal disparity coding in projections of medial superior olive to inferior colliculus. *J Neurosci* **23**, 7438-7449.

Oliver, D. L. & Huerta, M. F. (1992). Inferior and Superior Colliculi. In *The Mammalian Auditory Pathway: Neuroanatomy*, eds. Webster, D. B., Fay, R. R., & Popper, A. N., pp. 168-221. Springer-Verlag, New York.

Overholt, E. M., Rubel, E. W., & Hyson, R. L. (1992). A circuit for coding interaural time differences in the chick brainstem. *J Neurosci* **12**, 1698-1708.

Palmer, A. R. (1982). Encoding of rapid amplitude fluctuations by Cochlear-nerve fibres in the guinea-pig. *Arch Otorhinolaryngol* **236**, 197-202.

Palmer, A. R. & King, A. J. (1982). The representation of auditory space in the mammalian superior colliculus. *Nature* **299**, 248-249.

Palmer, A. R., Rees, A., & Caird, D. (1990). Interaural delay sensitivity to tones and broad band signals in the guinea-pig inferior colliculus. *Hear Res* **50**, 71-86.

Palmer, A. R. & Russell, I. J. (1986). Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells. *Hear Res* **24**, 1-15.

Park, T. J. (1998). IID sensitivity differs between two principal centers in the interaural intensity difference pathway: the LSO and the IC. *J Neurophysiol* **79**, 2416-2431.

Park, T. J., Grothe, B., Pollak, G. D., Schuller, G., & Koch, U. (1996). Neural delays shape selectivity to interaural intensity differences in the lateral superior olive. *J Neurosci* **16**, 6554-6566.

Park, T. J., Klug, A., Holinstat, M., & Grothe, B. (2004). Interaural level difference processing in the lateral superior olive and the inferior colliculus. *J Neurophysiol* **92**, 289-301.

Park, T. J. & Pollak, G. D. (1993). GABA shapes sensitivity to interaural intensity disparities in the mustache bat's inferior colliculus: implications for encoding sound location. *J Neurosci* **13**, 2050-2067.

Patterson, R. D. (1994). The sound of a sinusoid: Spectral models. *J Acoust Soc Am* **96**, 1409-1418.

Peruzzi, D., Sivaramakrishnan, S., & Oliver, D. L. (2000). Identification of cell types in brain slices of the inferior colliculus. *Neuroscience* **101**, 403-416.

Plenge, G. (1974). On the differences between localization and lateralization. *J Acoust Soc Am* **56**, 944-951.

Pollak, G. D., Burger, R. M., Park, T. J., Klug, A., & Bauer, E. E. (2002). Roles of inhibition for transforming binaural properties in the brainstem auditory system. *Hear Res* **168**, 60-78.

Popper, A. N. & Fay, R. R. (Eds) (1992). *The mammalian auditory pathway: Neurophysiology* Springer-Verlag, New York.

Rakerd, B. & Hartmann, W. M. (1986). Localization of sound in rooms, III: Onset and duration effects. *J Acoust Soc Am* **80**, 1695-1706.

Ramachandran R., Davis K.A. & May B. J. (1999). Single-unit responses in the inferior colliculus of decerebrate cats. I. Classification based on frequency response maps. J Neurophysiol **82**, 152-63.

Recanzone, G. H. & Beckerman, N. S. (2004). Effects of intensity and location on sound location discrimination in macaque monkeys. *Hear Res* **198**, 116-124.

Recio, A., Rich, N. C., Narayan, S. S., & Ruggero, M. A. (1998). Basilar-membrane responses to clicks at the base of the chinchilla cochlea. *J Acoust Soc Am* **103**, 1972-1989.

Rhode, W. S. (1971). Observations of the vibration of the basilar membrane in squirrel monkeys using the Mossbauer technique. *J Acoust Soc Am* **49**, Suppl.

Rhode, W. S. & Greenberg, S. (1992). Physiology of the Cochlear Nuclei. In *The Mammalian Auditory Pathway: Neurophysiology*, eds. Popper, A. N. & Fay, R. R., pp. 94-152. Springer-Verlag, New York.

Robinson, D. E. & Jeffress, L. A. (1963). Effect of varying the interaural noise correlation on the detectability of tonal signals. *J Acoust Soc Am* **35**, 1947-1952.

Robles, L. & Ruggero, M. A. (2001). Mechanics of the mammalian cochlea. *Physiol Rev* **81**, 1305-1352.

Rose, J. E., Brugge, J. F., Anderson, D. J., & Hind, J. E. (1967). Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *J Neurophysiol* **30**, 769-793.

Rose, J. E., Gross, N. B., Geisler, C. D., & Hind, J. E. (1966). Some neural mechanisms in the inferior colliculus of the cat which may be relevant to localization of a sound source. *J Neurophysiol* **29**, 288-314.

Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philos Trans R Soc Lond B Biol Sci* **336**, 367-373.

Rothman, J. S. & Manis, P. B. (2003). The roles potassium currents play in regulating the electrical activity of ventral cochlear nucleus neurons. *J Neurophysiol* **89**, 3097-3113.

Rothman, J. S., Young, E. D., & Manis, P. B. (1993). Convergence of auditory nerve fibers onto bushy cells in the ventral cochlear nucleus: implications of a computational model. *J Neurophysiol* **70**, 2562-2583.

Rowan, D. & Lutman, M. (2005). Generalisation of learning with ITD discrimination across frequency and type of cue. Assoc. Res. Otolaryngol. Abs.: 59.

Ruggero, M. A. (1992). Physiology and Coding of Sound in the Auditory Nerve. In *The Mammalian Auditory Pathway: Neurophysiology*, eds. Popper, A. N. & Fay, R. R., pp. 34-93. Springer-Verlag, New York.

Ruggero, M. A., Rich, N. C., Recio, A., Narayan, S. S., & Robles, L. (1997). Basilar-membrane responses to tones at the base of the chinchilla cochlea. *J Acoust Soc Am* **101**, 2151-2163.

Ryugo, D. K. (1992). The Auditory Nerve: Peripheral Innervation, Cell Body Morphology, and Central Projections. In *The Mammalian Auditory Pathway: Neuroanatomy*, eds. Webster, D. B., Fay, R. R., & Popper, A. N., pp. 23-65. Springer-Verlag, New York.

Saberi, K. & Perrott, D. R. (1990). Lateralization thresholds obtained under conditions in which the precedence effect is assumed to operate. *J Acoust Soc Am* **87**, 1732-1737.

Sally, S. L. & Kelly, J. B. (1992). Effects of superior olivary complex lesions on binaural responses in rat inferior colliculus. *Brain Res* **572**, 5-18.

Schreiner, C. E. & Langner, G. (1997). Laminar fine structure of frequency organization in auditory midbrain. *Nature* **388**, 383-386.

Scott, L. L., Mathews, P. J., & Golding, N. L. (2005). Post hearing developmental refinement of temporal processing in principal neurons of the medial superior olive. *J Neurosci* **25**, 7887-7895.

Seidl, A. H. & Grothe, B. (2005). Development of sound localization mechanisms in the Mongolian gerbil is shaped by early acoustic experience. *J Neurophysiol* **94**, 1028-1036.

Sellick, P. M., Patuzzi, R., & Johnstone, B. M. (1982). Measurement of basilar membrane motion in the guinea pig using the Mossbauer technique. *J Acoust Soc Am* **72**, 131-141.

Semple, M. N., Aitkin, L. M., Calford, M. B., Pettigrew, J. D., & Phillips, D. P. (1983). Spatial receptive fields in the cat inferior colliculus. *Hear Res* **10**, 203-215.

Senn, P., Kompis, M., Vischer, M., & Haeusler, R. (2005). Minimum audible angle, just noticeable interaural differences and speech intelligibility with bilateral cochlear implants using clinical speech processors. *Audiol Neurootol* **10**, 342-352.

Shackleton, T. M., McAlpine, D., & Palmer, A. R. (2000). Modelling convergent input onto interaural-delay-sensitive inferior colliculus neurones. *Hear Res* **149**, 199-215.

Shackleton, T. M. & Palmer, A. R. (2004). Sensitivity to changes in interaural time difference and interaural correlation in the inferior colliculus. In *Auditory signal processing: physiology, psychoacoustics, and models.* Springer Verlag, New York.

Shackleton, T. M., Skottun, B. C., Arnott, R. H., & Palmer, A. R. (2003). Interaural time difference discrimination thresholds for single neurons in the inferior colliculus of Guinea pigs. *J Neurosci* **23**, 716-724.

Shamma, S. A., Shen, N. M., & Gopalaswamy, P. (1989). Stereausis: binaural processing without neural delays. *J Acoust Soc Am* **86**, 989-1006.

Shaw, E. A. (1974). Transformation of sound pressure level from the free field to the eardrum in the horizontal plane. *J Acoust Soc Am* **56**, 1848-1861.

Shore, S. E., Sumner, C. J., Bledsoe, S. C., & Lu, J. (2003). Effects of contralateral sound stimulation on unit activity of ventral cochlear nucleus neurons. *Exp Brain Res* **153**, 427-435.

Simon, H. J., Collins, C. C., Jampolsky, A., Morledge, D. E., & Yu, J. (1994). The measurement of the lateralization of narrow bands of noise using an acoustic pointing paradigm: the effect of sound-pressure level. *J Acoust Soc Am* **95**, 1534-1547.

Skottun, B. C., Shackleton, T. M., Arnott, R. H., & Palmer, A. R. (2001). The ability of inferior colliculus neurons to signal differences in interaural delay. *Proc Natl Acad Sci U.S.A* **98**, 14050-14054.

Smith, P. H. (1995). Structural and functional differences distinguish principal from nonprincipal cells in the guinea pig MSO slice. *J Neurophysiol* **73**, 1653-1667.

Smith, P. H., Joris, P. X., Carney, L. H., & Yin, T. C. (1991). Projections of physiologically characterized globular bushy cell axons from the cochlear nucleus of the cat. *J Comp Neurol* **304**, 387-407.

Smith, P. H., Joris, P. X., & Yin, T. C. (1993). Projections of physiologically characterized spherical bushy cell axons from the cochlear nucleus of the cat: evidence for delay lines to the medial superior olive. *J Comp Neurol* **331**, 245-260.

Smith, P. H., Joris, P. X., & Yin, T. C. (1998). Anatomy and physiology of principal cells of the medial nucleus of the trapezoid body (MNTB) of the cat. *J Neurophysiol* **79**, 3127-3142.

Song, P., Yang, Y., Barnes-Davies, M., Bhattacharjee, A., Hamann, M., Forsythe, I. D., Oliver, D. L., & Kaczmarek, L. K. (2005). Acoustic environment determines phosphorylation state of the Kv3.1 potassium channel in auditory neurons. *Nat Neurosci* **8**, 1335-1342.

Speaks, C. (1996). *Introduction to Sound. Acoustics for the Hearing and Speech Sciences*, 2 ed. Singular Publishing Group Inc., San Diego, California.

Spitzer, M. W. & Semple, M. N. (1991). Interaural phase coding in auditory midbrain: influence of dynamic stimulus features. *Science* **254**, 721-724.

Spitzer, M. W. & Semple, M. N. (1995). Neurons sensitive to interaural phase disparity in gerbil superior olive: diverse monaural and temporal response properties. *J Neurophysiol* **73**, 1668-1690.

Stanford, T. R., Kuwada, S., & Batra, R. (1992). A comparison of the interaural time sensitivity of neurons in the inferior colliculus and thalamus of the unanesthetized rabbit. *J Neurosci* **12**, 3200-3216.

Stecker, G. C. (2005). Rate-limited, but accurate, central processing of interaural time differences in modulated high-frequency sounds. Focus on: "neural sensitivity to interaural envelope delays in the inferior colliculus of the guinea pig". *J Neurophysiol* **93**, 3048-3049.

Stecker, G. C. & Hafter, E. R. (2002). Temporal weighting in sound localization. *J Acoust Soc Am* **112**, 1046-1057.

Stecker, G. C., Harrington, I. A., Macpherson, E. A., & Middlebrooks, J. C. (2005a). Spatial sensitivity in the dorsal zone (area DZ) of cat auditory cortex. *J Neurophysiol*

Stecker, G. C., Harrington, I. A., & Middlebrooks, J. C. (2005b). Location coding by opponent neural populations in the auditory cortex. *PLoS Biol* **3**, e78.

Stecker, G. C., Mickey, B. J., Macpherson, E. A., & Middlebrooks, J. C. (2003). Spatial sensitivity in field PAF of cat auditory cortex. *J Neurophysiol* **89**, 2889-2903.

Stellmack, M. A., Viemeister, N. F., & Byrne, A. J. (2005). Discrimination of interaural phase differences in the envelopes of sinusoidally amplitude-modulated 4-kHz tones as a function of modulation depth. *J Acoust Soc Am* **118**, 346-352.

Sterbing, S. J., Hartung, K., & Hoffmann, K. P. (2003). Spatial tuning to virtual sounds in the inferior colliculus of the guinea pig. *J Neurophysiol* **90**, 2648-2659.

Stern, R. M. & Trahiotis, C. (1997). Models of binaural perception. In *Binaural and spatial hearing in real and virtual environments*, eds. Gilkey, R. H. & Anderson, T. R., pp. 499-532. Lawrence Erlbaum Associates, Inc., Mahwah, New Jersey.

Stern, R. M., Zeiberg, A. S., & Trahiotis, C. (1988). Lateralization of complex binaural stimuli: a weighted-image model. *J Acoust Soc Am* **84**, 156-165.

Stevens, S. S. & Newman, E. B. (1936). The Localization of actual sources of sound. *Am J Psychol* **48**, 297-306.

Strutt, J.W. (1907). On our perception of sound direction. Philos Mag 13, 214-232.
Sullivan, W. E. & Konishi, M. (1986). Neural map of interaural phase difference in the owl's brainstem. *Proc Natl Acad Sci U.S.A* **83**, 8400-8404.

Svirskis, G., Kotak, V., Sanes, D. H., & Rinzel, J. (2004). Sodium along with low-threshold potassium currents enhance coincidence detection of subthreshold noisy signals in MSO neurons. *J Neurophysiol* **91**, 2465-2473.

Tasaki, I. (1954). Nerve impulses in individual auditory nerve fibers of guinea pig. *J Neurophysiol* **17**, 97-122.

Tobias, J. V. & Schubert, E. D. (1959). Effective onset duration of auditory stimuli. *J Acoust Soc Am* **31**, 1595-1605.

Tobias, J. V. & Zerlin, S. (1959). Lateralization threshold as a function of duration. *J Acoust Soc Am* **31**, 1591-1594.

Tollin, D. J. (2003). The lateral superior olive: a functional role in sound source localization. *Neuroscientist* **9**, 127-143.

Tollin, D. J., Populin, L. C., Moore, J. M., Ruhland, J. L., & Yin, T. C. (2005). Sound-localization performance in the cat: the effect of restraining the head. *J Neurophysiol* **93**, 1223-1234.

Tollin, D. J., Populin, L. C., & Yin, T. C. (2004). Neural correlates of the precedence effect in the inferior colliculus of behaving cats. *J Neurophysiol* **92**, 3286-3297.

Tollin, D. J. & Yin, T. C. (2002a). The coding of spatial location by single units in the lateral superior olive of the cat. I. Spatial receptive fields in azimuth. *J Neurosci* **22**, 1454-1467.

Tollin, D. J. & Yin, T. C. (2002b). The coding of spatial location by single units in the lateral superior olive of the cat. II. The determinants of spatial receptive fields in azimuth. *J Neurosci* **22**, 1468-1479.

Tollin, D. J. & Yin, T. C. (2005). Interaural Phase and Level Difference Sensitivity in Low-Frequency Neurons in the Lateral Superior Olive. *J Neurosci* **25**, 10648-10657.

Trahiotis, C. & Bernstein, L. R. (1986). Lateralization of bands of noise and sinusoidally amplitude-modulated tones: effects of spectral locus and bandwidth. *J Acoust Soc Am* **79**, 1950-1957.

Trahiotis, C., Bernstein, L. R., Buell, T. N., & Spektor, Z. (1990). On the use of adaptive procedures in binaural experiments. *J Acoust Soc Am* **87**, 1359-1361.

Trahiotis, C. & Stern, R. M. (1989). Lateralization of bands of noise: effects of bandwidth and differences of interaural time and phase. *J Acoust Soc Am* **86**, 1285-1293.

Trussell, L. O. (1999). Synaptic mechanisms for coding timing in auditory neurons. *Annu Rev Physiol* **61**, 477-496.

Tsuchitani, C. & Boudreau, J. C. (1966). Single unit analysis of cat superior olive S segment with tonal stimuli. *J Neurophysiol* **29**, 684-697.

Tsuchitani, C. & Boudreau, J. C. (1969). Stimulus level of dichotically presented tones and cat superior olive S-segment cell discharge. *J Acoust Soc Am* **46**, 979-988.

van der, Heijden, M. & Trahiotis, C. (1999). Masking with interaurally delayed stimuli: the use of "internal" delays in binaural detection. *J Acoust Soc Am* **105**, 388-399.

van de Par, S. & Kohlrausch, A. (1995). Analytic expressions for envelope correlations of certain narrow-band stimuli. *J Acoust Soc Am* 98, 3157-3169.

van de Par, S. & Kohlrausch, A. (1997). A new approach to comparing binaural masking level differences at low and high frequencies. *J Acoust Soc Am* **101**, 1671-1680.

van Hoesel, R. J. & Tyler, R. S. (2003). Speech perception, localization, and lateralization with bilateral cochlear implants. *J Acoust Soc Am* **113**, 1617-1630.

Versnel, H., Schoonhoven, R., Prijs, V. F. (1992). Single-fibre and whole-nerve responses to clicks as a function of sound intensity in the guinea pig. *Hear Res* **59**, 138-56.

von Bekesy, G. (1947). The variation of phase along the basilar membrane with sinusoidal vibration. *J Acoust Soc Am* **19**, 452-460.

Wagner, H., Takahashi, T., & Konishi, M. (1987). Representation of interaural time difference in the central nucleus of the barn owl's inferior colliculus. *J Neurosci* **7**, 3105-3116.

Wallach, H., Newman, E. B., & Rosenzweig, R. (1949). The precedence effect in sound localization. *Am J Psychol* **32**, 315-336.

Wang, X., Lu, T., Snider, R. K., & Liang, L. (2005). Sustained firing in auditory cortex evoked by preferred stimuli. *Nature* **435**, 341-346.

Webster, D. B., Popper, A. N., & Fay, R. R. (Eds) (1992). *The mammalian auditory pathway: Neuroanatomy* Springer-Verlag, New York.

Webster, F. A. (1951). The influence of interaural phase on masked thresholds I. The role of interaural time-deviation. *J Acoust Soc Am* **23**, 452-462.

Weiss, T. F. & Rose, C. (1988). A comparison of synchronization filters in different auditory receptor organs. *Hear Res* **33**, 175-179.

Wenzel, E. M., Arruda, M., Kistler, D. J., & Wightman, F. L. (1993). Localization using non-individualized head-related transfer functions. *J Acoust Soc Am* **94**, 111-123.

Whitworth, R.H. & Jeffress, L. A. (1961). Time vs. intensity in the localization of Tones. *J Acoust Soc Am* **33**, 925-929.

Wightman, F. L. & Kistler, D. J. (1989a). Headphone simulation of free-field listening. I: Stimulus synthesis. *J Acoust Soc Am* **85**, 858-867.

Wightman, F. L. & Kistler, D. J. (1989b). Headphone simulation of free-field listening. II: Psychophysical validation. *J Acoust Soc Am* **85**, 868-878.

Wightman, F. L. & Kistler, D. J. (1992). The dominant role of low-frequency interaural time differences in sound localization. *J Acoust Soc Am* **91**, 1648-1661.

Wightman, F. L. & Kistler, D. J. (1997). Monaural sound localization revisited. *J Acoust Soc Am* **101**, 1050-1063.

Winer, J. A. (1992). The Functional Architecture of the Medial Geniculate Body and the Primary Auditory Cortex. In *The Mammalian Auditory Pathway: Neuroanatomy*, eds. Webster, D. B., Popper, A. N., & Fay, R. R., pp. 222-409. Springer-Verlag, New York.

Wu, S. H. & Kelly, J. B. (1992). Binaural interaction in the lateral superior olive: time difference sensitivity studied in mouse brain slice. *J Neurophysiol* **68**, 1151-1159.

Yates, G. K., Robertson, D., & Johnstone, B. M. (1985). Very rapid adaptation in the guinea pig auditory nerve. *Hear Res* **17**, 1-12.

Yin, T. C. (2001). Neural mechanisms of encoding binaural localization cues. In *Integrative functions in the mammalian auditory pathway*, eds. Oertel, D., Fay, R. R., & Popper, A. N., pp. 99-159. Springer, New York.

Yin, T. C. & Chan, J. C. (1990). Interaural time sensitivity in medial superior olive of cat. *J Neurophysiol* **64**, 465-488.

Yin, T. C., Chan, J. C., & Carney, L. H. (1987). Effects of interaural time delays of noise stimuli on low-frequency cells in the cat's inferior colliculus. III. Evidence for cross-correlation. *J Neurophysiol* **58**, 562-583.

Yin, T. C., Chan, J. C., & Irvine, D. R. (1986). Effects of interaural time delays of noise stimuli on low-frequency cells in the cat's inferior colliculus. I. Responses to wideband noise. *J Neurophysiol* **55**, 280-300.

Yin, T. C. & Kuwada, S. (1983a). Binaural interaction in low-frequency neurons in inferior colliculus of the cat. II. Effects of changing rate and direction of interaural phase. *J Neurophysiol* **50**, 1000-1019.

Yin, T. C. & Kuwada, S. (1983b). Binaural interaction in low-frequency neurons in inferior colliculus of the cat. III. Effects of changing frequency. *J Neurophysiol* **50**, 1020-1042.

Yin, T. C., Kuwada, S., & Sujaku, Y. (1984). Interaural time sensitivity of high-frequency neurons in the inferior colliculus. *J Acoust Soc Am* **76**, 1401-1410.

Yost, W. A. (1974). Discriminations of interaural phase differences. *J Acoust Soc Am* **55**, 1299-1303.

Yost, W. A. & Soderquist, D. R. (1984). The precedence effect: revisited. *J Acoust Soc Am* **76**, 1377-1383.

Yost, W. A., Wightman, F. L., & Green, D. M. (1971). Lateralization of filtered clicks. *J Acoust Soc Am* **50**, 1526-1531.

Young, E. D. & Davies, K. A. (2001). Circuitry and function of the dorsal cohlear nucleus. In *Integrative functions in the mammalian auditory pathway*, eds. Oertel, D., Fay, R. R., & Popper, A. N., pp. 160-206. Springer, New York.

Young, S. R. & Rubel, E. W. (1983). Frequency-specific projections of individual neurons in chick brainstem auditory nuclei. *J Neurosci* **3**, 1373-1378.

Zacksenhouse, M., Johnson, D. H., & Tsuchitani, C. (1995). Transient effects during the chopping response of LSO neurons. *J Acoust Soc Am* **98**, 1410-1422.

Zatorre, R. J. & Penhune, V. B. (2001). Spatial localization after excision of human auditory cortex. *J Neurosci* **21**, 6321-6328.

Zhou, Y., Carney, L. H., & Colburn, H. S. (2005). A model for interaural time difference sensitivity in the medial superior olive: interaction of excitatory and inhibitory synaptic inputs, channel dynamics, and cellular morphology. *J Neurosci* **25**, 3046-3058.

Zurek, P. M. (1980). The precedence effect and its possible role in the avoidance of interaural ambiguities. *J Acoust Soc Am* **67**, 953-964.

Zurek, P. M. (1993). A note on onset effects in binaural hearing. *J Acoust Soc Am* **93**, 1200-1201.

Zwislocki, J. & Feldman, R. S. (1956). Just noticeable differences in dichotic phase. *J Acoust Soc Am* **28**, 860-864.